

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Takashi OKADO et al.
- (ii) TITLE OF INVENTION: GENE REGULATING AUREOBASIDIN SENSITIVITY
- (iii) NUMBER OF SEQUENCES: 50
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: WENDEROTH, LIND & PONACK, L.L.P.
 - (B) STREET: 2033 K Street, N.W., #800
 - (C) CITY: Washington
 - (D) STATE: D.C.
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 20006
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: MS-DOS
 - (D) SOFTWARE: Wordperfect 5.1
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/951,217
 - (B) FILING DATE: September 12, 2001
 - (C) CLASSIFICATION:

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 09/368,431

(B) FILING DATE: August 5, 1999

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/716,873

(B) FILING DATE: September 20, 1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Warren M. Cheek, Jr..

(B) REGISTRATION NUMBER: 33,367

(C) REFERENCE/DOCKET NUMBER:

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202-721-8200

(B) TELEFAX: 202-721-8250

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4140 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

AGATCTGTGG	CTTCCGGTTG	GCTACTTGTA	ACCAACTGAT	GGTCAGATGG	ATCTGCCGTC	60
TGTTTTGATT	TGAATTTTCC	CTGCTCATTC	TGATTCTGTG	AGAGGCTGCA	TTCATTATCA	120
CATCTCATA	CCGGCGCCTG	CGACTTCGGT	CACCTCTGCG	GTCTGGCGGT	TACCGGGGTC	180
CGTCTGAGAC	TCGTCAGTCA	GCCATTCGAG	TATGCGAACT	CTGACTTTGC	TCACCTAAGA	240
GTTTGCACGA	GATGCCGAAA	TCCTCCTCGA	GTAGAGTTTG	CAAGGCTTGA	ACCTTGGTCC	300
TTGAAGCCCG	AAAGTGGCTC	AGTAGTGGGA	TCGATAGTCT	GGTTGTTGAA	GATTTTCTCC	360
TCCACCTTAC	CTATGGCCGC	TGGCCTTCTC	CACCTTTCAG	GCTTTCAGGC	ACCCTCGGCT	420
CGGATTCTGT	ATCGTCCGGT	ACCGAAGCTA	GTCCTAGCTA	GTCAAAGCTA	GTCCAAGCTA	480
GTCTCGTCAA	GGTTTGGCGC	AGCGCGGTTC	CGTGTAAGT	ACAAATTTGA	AATACGAATA	540
CGCAGTACTC	GCAGCCGGCA	CTTCCGCTCA	GCCCAGGCTC	AGAGGCTAAG	GGTGTTGGCG	600
CTTCCTCATC	ATCTTCTTCT	CGTCGACCTT	TTCCTCTTTC	TCTCCCTATC	GGTGCTTCTC	660
TCCAACCTCA	TTCTCAGTCG	TTCGCCCATC	AGGTTTATAC	TCCGGCTCCG	TGGCCATCTG	720
CCTCCCTCAC	GACCTCCTCG	TTCCAGGTTT	TCCTCTCGAC	TGCTGCGCCC	TTGCACTTCG	780
CCTTGCAATCA	GTGAAACCCC	CTGCAACGTG	ACGGCTCAAA	GACATCCTCG	TTTGGCCGCT	840
GGAGACCGGA	GCGTGCGCTT	CGTTTCGTCT	TCTTCGAACC	GATCTCAATT	TCCCCGCTCG	900
GGTTGACGCC	GTCAGCACCC	TGCTCGTTGC	CTAACGGCTT	GTTATTCAAG	ACCCCTTTTC	960
TGCCGCTTCC	GCGACCGATT	TATTCGTCGC	CTTCCAATC	TTGTACAATC	GGGGGGAAAG	1020
AAAGCAGACG	GAGTTCGATC	TGGAGGAATT	ATAGCTGAGT	CTTGCCCGCA	AGACTCGCCG	1080
CAACCATGAA	TCAAACACTT	CCCACGTGGA	AGGACCGCAC	GCAGAACCAG	TTTGGAAGC	1140
TTCAGATCCA	GGTTCCATGG	CGGTCCATCC	AACTGCTCGT	CCCGCATCGC	ATGCGGCGGA	1200
AGTTAAGGTC	CAAATTGCGC	AGTAGAGCGT	CTCCTACCTC	GTCAATAGCC	TCTTTACAGA	1260
CGTCGTTATC	GCCTGCAGAC	AACTACGAT	CGCTCCAAAG	CCACCGATGG	ACGGTTTACG	1320

ACTTCCAATA TCTGCTTCTG TTGATCGTGG GCATCTTCTC TTTGACCGTT ATCGAGTCGC 1380
 CCGGGCCTTT GGGCAAAACG GCCATTTTCT CCATGCTCCT ATTCTCTCTC CTGATCCCTA 1440
 TGACCCGCCA GTTCTTCCTC CCGTTTCTGC CGATTGCCGG ATGGCTTCTG TTTTCTACG 1500
 CCTGCCAGTG AGTTAAAAAC AACCCGCTAC CAGACCCCGT GCAGCAGTTA CTCACATATG 1560
 CAGGTTCATC CCAAGCGATT GGCGCCCTGC GATTTGGGTT CGTGTCTTGC CTGCACTGGA 1620
 GAATATTCTC TACGGCGCAA ACATCAGCAA CATCCTATCC GCTCACCAGA ACGTTGTGCT 1680
 TGACGTGCTG GCGTGGCTAC CCTACGGTAT CTGCCACTAT GGCGCTCCGT TTGTGTGCTC 1740
 GTTGATCATG TTCATCTTCG GTCCGCCCCG CACTGTTCCC CTTTTCGCGC GCACTTTCGG 1800
 CTATATCAGT ATGACTGCGG TTACTATTCA GCTGTTTTTC CCTTGCTCTC CACCTTGGTA 1860
 TGAGAATCGC TATGGTCTAG CTCCGGCAGA CTACTCCATC CAAGGTGATC CCGCAGGGCT 1920
 TGCCCGCATT GACAAGCTTT TCGGCATCGA CCTTTACACG TCTGTTTTCC ATCAGTCGCC 1980
 TGTGTGTTC GGCGCTTTTC CGTCGCTGCA TGCTGCCGAC TCAACCCTGG CCGCACTTTT 2040
 CATGAGTCAT GTTTTCCCCC GCATGAAGCC CGTCTTCGTG ACCTATACTC TATGGATGTG 2100
 GTGGGCAACA ATGTACCTCT CACATCACTA TGCGGTCGAT TTGGTTGCGG GTGGTCTCCT 2160
 GGCCGCCATT GCTTTCTACT TCGCCAAGAC CCGATTCCCTT CCCCCTGTCC AGCTCGACAA 2220
 GACCTTCCGT TGGGACTACG ACTATGTGGA ATTCGGCGAG TCTGCCCTGG AGTATGGGTA 2280
 TGGTGCAGCT GGCTATGATG GAGACTTCAA TCTCGACAGC GATGAATGGA CTGTTGGTTC 2340
 TTCATCCTCC GTCTCCTCAG GCTCCTTGAG TCCCGTTGAC GATCATTACT CATGGGAAAC 2400
 CGAGGCACTG ACCTCCCCAC ATACTGATAT TGAGTCCGGC AGGCATACTT TCAGCCCTTG 2460
 AGTAGCCACA AACCAAACCTC GATACCTGCA TATAGCGATC TCGCTCCTCC TCCACTGCAT 2520
 CTATTTACGA GACGGCGTTA GAACATTTCA CGACATTCTG GCTTTATTGC ATCGAGCACA 2580
 TTTTCGACACA TATATCTTTA ATACCCTTTC TTCGGTGTCC CAGATCATCG GTTCGACCTT 2640
 AATGTACCTC GGTCCGAATC CGCCTGGGAT ACTGTTTCTC TTTCCGCCGC ACTTCACTGT 2700

ACATTGCTTG ACATTGCGAA ACCGGGTTGG GCTCGAACGT GGGATGGGTT ATCGCTCATC 2760
GCTACACGCC GTTGCTCCAT CATAATGTTA ATGGACACAA TGGGGCTACG CATCCTGGTG 2820
TTTAGTCCTG GAAGACCATC CGATAACCCC CGTCGGTAAC ACTCGCTTGT CTCGTGTCCA 2880
CCCAGACACT ACTTCAATTC TCACTTCTAT CGTCCGCTAT TACCTTGACC TGGTCGAACC 2940
CATCCTTATT ATTCGTTTCG ACTATGCTAT ATATTTATTT TTACCATTCTG TGTCGATCGC 3000
TCATACTCTT GGCCTTGGG ACTGGAAGCA TTTATATTGG AAAAAATCAC GGAATGGGGC 3060
GCCTTTTCTT CTTGCACTTC ACTCGCTGTG CATAGACGGT TTTACATTTC TGCTTTGCAA 3120
TGCATCACGA ACTCTGCATT AGCATATAGA AAGAGGGGAA GGATGGACCT TCTTCTTGAT 3180
TGCTCGCATG GTTTATCCAT TCGCTCAAAG TGGATTACGT CCACATATTA CCCGGGGGCT 3240
ATACACATGG CTA CTGTGT GCTTTCTGAC ATTCGCCGGA CGTGCAAGGT TGGGAGGAGA 3300
GTCTGACGCT GACGGGGCTT GTTGAAGGAT GTTCACGCGT CCCGATTGTA CCCGGCTTCG 3360
ACTAACCTCA GATTCTCGAC TTGTTGGACG GTGACTTGAC TTGCTTGCTA TGGTCTGACG 3420
CTCTCACACC TACCTATCAC ATCCTCCTCA CCTCACAAAT TCCGCTCATG GACACTATCC 3480
TCTTCTTTTC GTTTCCTTG GATAGTGTGT GTGTGTGTGT GGTGGGGCA AATTATCCAT 3540
AGCAGCAGTA TTATTAGTTA TAATCCGGTA GTGTTATGAT TTATGAAGGC AACTTGTATA 3600
CTATTGCCAC TTTGTCCATA TCTCTTGCTT GTAATAGAAC TGACATCGCG ACGCTTCGGC 3660
CACGATGCAT ATAAAACTC TAGTCAACAC GATATTAACA AGCGAAACCA TTACGCTGTA 3720
AACTATTCAG GATCGCCGCG GGCCCATCTG GGACTTGACT GTACTAAATA TGTCCTAAAG 3780
CAAGCAGACT AAATATTTAA CGTGGGATAT TATTCATATA CGCATATGTA TACATAGTCA 3840
TAACAAGCCA AGGGGTGGGT AGGGGTGGGT AATTATTATT TTTTTCGTC GATACAAGTA 3900
TCCATCCTTA AATGTCCGTG GTCTACTCTT CATAAATCTT AACCCGCTCC GCATACTCCT 3960
TTATCCTCGA GACAAAAGTG TCTTCAATTT CATCGCCACG GCCACCAGCA ACGCGGAGGA 4020
TAAGGTCGTT GAGAGAGGCG CGGCCGAGAA TGTCGACATG GTCGCCTTCA TATTGTTAGC 4080

ATGCGACGTC AGACTGGAAC CAGGAAGGGA AAGGAGAGAG GTACCTGTAT TTGGACCACC 4140

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 439 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asn	Gln	Thr	Leu	Pro	Thr	Trp	Lys	Asp	Arg	Thr	Gln	Asn	Gln
1				5					10					15
Phe	Gly	Lys	Leu	Gln	Ile	Gln	Val	Pro	Trp	Arg	Ser	Ile	Gln	Leu
				20					25					30
Leu	Val	Pro	His	Arg	Met	Arg	Arg	Lys	Leu	Arg	Ser	Lys	Leu	Arg
				35					40					45
Ser	Arg	Ala	Ser	Pro	Thr	Ser	Ser	Ile	Ala	Ser	Leu	Gln	Thr	Ser
				50					55					60
Leu	Ser	Pro	Ala	Asp	Thr	Leu	Arg	Ser	Leu	Gln	Ser	His	Arg	Trp
				65					70					75
Thr	Val	Tyr	Asp	Phe	Gln	Tyr	Leu	Leu	Leu	Leu	Ile	Val	Gly	Ile
				80					85					90

Phe	Ser	Leu	Thr	Val	Ile	Glu	Ser	Pro	Gly	Pro	Leu	Gly	Lys	Thr
				95					100					105
Ala	Ile	Phe	Ser	Met	Leu	Leu	Phe	Ser	Leu	Leu	Ile	Pro	Met	Thr
				110					115					120
Arg	Gln	Phe	Phe	Leu	Pro	Phe	Leu	Pro	Ile	Ala	Gly	Trp	Leu	Leu
				125					130					135
Phe	Phe	Tyr	Ala	Cys	Gln	Phe	Ile	Pro	Ser	Asp	Trp	Arg	Pro	Ala
				140					145					150
Ile	Trp	Val	Arg	Val	Leu	Pro	Ala	Leu	Glu	Asn	Ile	Leu	Tyr	Gly
				155					160					165
Ala	Asn	Ile	Ser	Asn	Ile	Leu	Ser	Ala	His	Gln	Asn	Val	Val	Leu
				170					175					180
Asp	Val	Leu	Ala	Trp	Leu	Pro	Tyr	Gly	Ile	Cys	His	Tyr	Gly	Ala
				185					190					195
Pro	Phe	Val	Cys	Ser	Leu	Ile	Met	Phe	Ile	Phe	Gly	Pro	Pro	Gly
				200					205					210
Thr	Val	Pro	Leu	Phe	Ala	Arg	Thr	Phe	Gly	Tyr	Ile	Ser	Met	Thr
				215					220					225
Ala	Val	Thr	Ile	Gln	Leu	Phe	Phe	Pro	Cys	Ser	Pro	Pro	Trp	Tyr
				230					235					240
Glu	Asn	Arg	Tyr	Gly	Leu	Ala	Pro	Ala	Asp	Tyr	Ser	Ile	Gln	Gly
				245					250					255
Asp	Pro	Ala	Gly	Leu	Ala	Arg	Ile	Asp	Lys	Leu	Phe	Gly	Ile	Asp

260	265	270
Leu Tyr Thr Ser Val Phe His Gln Ser Pro Val Val Phe Gly Ala		
275	280	285
Phe Pro Ser Leu His Ala Ala Asp Ser Thr Leu Ala Ala Leu Phe		
290	295	300
Met Ser His Val Phe Pro Arg Met Lys Pro Val Phe Val Thr Tyr		
305	310	315
Thr Leu Trp Met Trp Trp Ala Thr Met Tyr Leu Ser His His Tyr		
320	325	330
Ala Val Asp Leu Val Ala Gly Gly Leu Leu Ala Ala Ile Ala Phe		
335	340	345
Tyr Phe Ala Lys Thr Arg Phe Leu Pro Arg Val Gln Leu Asp Lys		
350	355	360
Thr Phe Arg Trp Asp Tyr Asp Tyr Val Glu Phe Gly Glu Ser Ala		
365	370	375
Leu Glu Tyr Gly Tyr Gly Ala Ala Gly Tyr Asp Gly Asp Phe Asn		
380	385	390
Leu Asp Ser Asp Glu Trp Thr Val Gly Ser Ser Ser Ser Val Ser		
395	400	405
Ser Gly Ser Leu Ser Pro Val Asp Asp His Tyr Ser Trp Glu Thr		
410	415	420
Glu Ala Leu Thr Ser Pro His Thr Asp Ile Glu Ser Gly Arg His		
425	430	435

Thr Phe Ser Pro

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2856 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GGTTTATACT CCGGCTCCGT GGCCATCTGC CTCCCTCACG ACCTCCTCGT TCCAGGTTTT 60
CCTCTCGACT GCTGCGCCCT TGCACTTCGC CTTGCATCAG TGAAACCCCC TGCAACGTGA 120
CGGCTCAAAG ACATCCTCGT TTGGCCGCTG GAGACCGGAG CGTGCGCTTC GTTTCGTCTT 180
CTTCGAACCG ATCTCAATTT CCCCCTCTGG GTTGACGCCG TCAGCACCTT GCTCGTTGCC 240
TAACGGCTTG TTATTCAAGA CCCCTTTTCT GCCGCTTCCG CGACCGATTT ATTCGTCGCC 300
TTCCAACCTCT TGTACAATCG GGGGGAAAGA AAGCAGACGG AGTTCGATCT GGAGGAATTA 360
TAGCTGAGTC TTGCCCCGAA GACTCGCCGC AACCATGAAT CAAACACTTC CCACGTGGAA 420
GGACCGCACG CAGAACCAGT TTGGAAAGCT TCAGATCCAG GTTCCATGGC GTTCCATCCA 480
ACTGCTCGTC CCGCATCGCA TGCGGCGGAA GTTAAGGTCC AAATTGCGCA GTAGAGCGTC 540
TCCTACCTCG TCAATAGCCT CTTTACAGAC GTCGTTATCG CCTGCAGACA CACTACGATC 600
GCTCCAAAGC CACCGATGGA CGGTTTACGA CTTCCAATAT CTGCTTCTGT TGATCGTGGG 660
```

CATCTTCTCT	TTGACCGTTA	TCGAGTCGCC	CGGGCCTTTG	GGCAAAACGG	CCATTTTCTC	720
CATGCTCCTA	TTCTCTCTCC	TGATCCCTAT	GACCCGCCAG	TTCTTCCTCC	CGTTTCTGCC	780
GATTGCCGGA	TGGCTTCTGT	TTTTCTACGC	CTGCCAGTTC	ATCCCAAGCG	ATTGGCGCCC	840
TGCGATTTGG	GTTCGTGTCT	TGCCTGCACT	GGAGAATATT	CTCTACGGCG	CAAACATCAG	900
CAACATCCTA	TCCGCTCACC	AGAACGTTGT	GCTTGACGTG	CTGGCGTGGC	TACCCTACGG	960
TATCTGCCAC	TATGGCGCTC	CGTTTGTGTG	CTCGTTGATC	ATGTTCATCT	TCGGTCCGCC	1020
CGGCACTGTT	CCCCTTTTCG	CGCGCACTTT	CGGCTATATC	AGTATGACTG	CGGTTACTAT	1080
TCAGCTGTTT	TTCCCTTGCT	CTCCACCTTG	GTATGAGAAT	CGCTATGGTC	TAGCTCCGGC	1140
AGACTACTCC	ATCCAAGGTG	ATCCCGCAGG	GCTTGCCCGC	ATTGACAAGC	TTTTCGGCAT	1200
CGACCTTTAC	ACGTCTGTTT	TCCATCAGTC	GCCTGTTGTG	TTCGGCGCTT	TTCCGTCGCT	1260
GCATGCTGCC	GACTCAACCC	TGGCCGCACT	TTTCATGAGT	CATGTTTTCC	CCCGCATGAA	1320
GCCCGTCTTC	GTGACCTATA	CTCTATGGAT	GTGGTGGGCA	ACAATGTACC	TCTCACATCA	1380
CTATGCGGTC	GATTTGGTTG	CGGGTGGTCT	CCTGGCCGCC	ATTGCTTTCT	ACTTCGCCAA	1440
GACCCGATTC	CTTCCCCGTG	TCCAGCTCGA	CAAGACCTTC	CGTTGGGACT	ACGACTATGT	1500
GGAATTCGGC	GAGTCTGCCC	TGGAGTATGG	GTATGGTGCA	GCTGGCTATG	ATGGAGACTT	1560
CAATCTCGAC	AGCGATGAAT	GGACTGTTGG	TTCTTCATCC	TCCGTCTCCT	CAGGCTCCTT	1620
GAGTCCCGTT	GACGATCATT	ACTCATGGGA	AACCGAGGCA	CTGACCTCCC	CACATACTGA	1680
TATTGAGTCC	GGCAGGCATA	CTTTCAGCCC	TTGAGTAGCC	ACAAACCAA	CTCGATACCT	1740
GCATATAGCG	ATCTCGCTCC	TCCTCCACTG	CATCTATTTA	CGAGACGGCG	TTAGAACATT	1800
TCACGACATT	CTGGCTTTAT	TGCATCGAGC	ACATTTCGAC	ACATATATCT	TTAATACCCT	1860
TTCTTCGGTG	TCCCAGATCA	TCGGTTCGAC	CTTAATGTAC	CTCGGTCCGA	ATCCGCCTGG	1920
GATACTGTTT	CTCTTTCCGC	CGCACTTCAC	TGTACATTGC	TTGACATTGC	GAAACCGGGT	1980
TGGGCTCGAA	CGTGGGATGG	GTTATCGCTC	ATCGCTACAC	GCCGTTGCTC	CATCATAATG	2040

TTAATGGACA CAATGGGGCT ACGCATCCTG GTGTTTAGTC CTGGAAGACC ATCCGATAAC 2100
 CCCCCTCGGT AACACTCGCT TGTCTCGTGT CCACCCAGAC ACTACTTCAA TTCTCACTTC 2160
 TATCGTCCGC TATTACCTTG ACCTGGTCGA ACCCATCCTT ATTATTCGTT TCGACTATGC 2220
 TATATATTTA TTTTACCAT TCGTGTCGAT CGCTCATACT CTTGGCGCTT GGGACTGGAA 2280
 GCATTTATAT TGGAAAAAAT CACGGAATGG GGCGCCTTTT CTTCTTGAC TCACTCGCT 2340
 GTGCATAGAC GGTTTTACAT TTCTGCTTTG CAATGCATCA CGAACTCTGC ATTAGCATAT 2400
 AGAAAGAGGG GAAGGATGGA CCTTCTTCTT GATTGCTCGC ATGGTTTATC CATTCGCTCA 2460
 AAGTGGATTA CGTCCACATA TTACCCGGGG GCTATACACA TGGCTACTGT GTTGCTTTCT 2520
 GACATTCGCC GGACGTGCAA GGTGGGAGG AGAGTCTGAC GCTGACGGGG CTTGTTGAAG 2580
 GATGTTCACG CGTCCCGATT TGACCCGGCT TCGACTAACC TCAGATTCTC GACTTGTTGG 2640
 ACGGTGACTT GACTTGCTTG CTATGGTCTG ACGCTCTCAC ACCTACCTAT CACATCCTCC 2700
 TCACCTCACA AATTCCGCTC ATGGACACTA TCCTCTTCTT TTCGTTTCCC TTGGATAGTG 2760
 TGTGTGTGTG TGTGGTTGGG GCAAATTATC CATAGCAGCA GTATTATTAG TTATAATCCG 2820
 GTAGTGTTAT GATTTATGAA GGCAACTTGT ATACTA 2856

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 439 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asn	Gln	Thr	Leu	Pro	Thr	Trp	Lys	Asp	Arg	Thr	Gln	Asn	Gln
1				5					10					15
Phe	Gly	Lys	Leu	Gln	Ile	Gln	Val	Pro	Trp	Arg	Ser	Ile	Gln	Leu
				20					25					30
Leu	Val	Pro	His	Arg	Met	Arg	Arg	Lys	Leu	Arg	Ser	Lys	Leu	Arg
				35					40					45
Ser	Arg	Ala	Ser	Pro	Thr	Ser	Ser	Ile	Ala	Ser	Leu	Gln	Thr	Ser
				50					55					60
Leu	Ser	Pro	Ala	Asp	Thr	Leu	Arg	Ser	Leu	Gln	Ser	His	Arg	Trp
				65					70					75
Thr	Val	Tyr	Asp	Phe	Gln	Tyr	Leu	Leu	Leu	Leu	Ile	Val	Gly	Ile
				80					85					90
Phe	Ser	Leu	Thr	Val	Ile	Glu	Ser	Pro	Gly	Pro	Leu	Gly	Lys	Thr
				95					100					105
Ala	Ile	Phe	Ser	Met	Leu	Leu	Phe	Ser	Leu	Leu	Ile	Pro	Met	Thr
				110					115					120
Arg	Gln	Phe	Phe	Leu	Pro	Phe	Leu	Pro	Ile	Ala	Gly	Trp	Leu	Leu
				125					130					135
Phe	Phe	Tyr	Ala	Cys	Gln	Phe	Ile	Pro	Ser	Asp	Trp	Arg	Pro	Ala
				140					145					150
Ile	Trp	Val	Arg	Val	Leu	Pro	Ala	Leu	Glu	Asn	Ile	Leu	Tyr	Gly
				155					160					165
Ala	Asn	Ile	Ser	Asn	Ile	Leu	Ser	Ala	His	Gln	Asn	Val	Val	Leu

170	175	180
Asp Val Leu Ala Trp Leu Pro Tyr Gly Ile Cys His Tyr Gly Ala		
185	190	195
Pro Phe Val Cys Ser Leu Ile Met Phe Ile Phe Gly Pro Pro Gly		
200	205	210
Thr Val Pro Leu Phe Ala Arg Thr Phe Gly Tyr Ile Ser Met Thr		
215	220	225
Ala Val Thr Ile Gln Leu Phe Phe Pro Cys Ser Pro Pro Trp Tyr		
230	235	240
Glu Asn Arg Tyr Gly Leu Ala Pro Ala Asp Tyr Ser Ile Gln Gly		
245	250	255
Asp Pro Ala Gly Leu Ala Arg Ile Asp Lys Leu Phe Gly Ile Asp		
260	265	270
Leu Tyr Thr Ser Gly Phe His Gln Ser Pro Val Val Phe Gly Ala		
275	280	285
Phe Pro Ser Leu His Ala Ala Asp Ser Thr Leu Ala Ala Leu Phe		
290	295	300
Met Ser His Val Phe Pro Arg Met Lys Pro Val Phe Val Thr Tyr		
305	310	315
Thr Leu Trp Met Trp Trp Ala Thr Met Tyr Leu Ser His His Tyr		
320	325	330
Ala Val Asp Leu Val Ala Gly Gly Leu Leu Ala Ala Ile Ala Phe		
335	340	345

Tyr Phe Ala Lys Thr Arg Phe Leu Pro Arg Val Gln Leu Asp Lys		
	350	355 360
Thr Phe Arg Trp Asp Tyr Asp Tyr Val Glu Phe Gly Glu Ser Ala		
	365	370 375
Leu Glu Tyr Gly Tyr Gly Ala Ala Gly Tyr Asp Gly Asp Phe Asn		
	380	385 390
Leu Asp Ser Asp Glu Trp Thr Val Gly Ser Ser Ser Ser Val Ser		
	395	400 405
Ser Gly Ser Leu Ser Pro Val Asp Asp His Tyr Ser Trp Glu Thr		
	410	415 420
Glu Ala Leu Thr Ser Pro His Thr Asp Ile Glu Ser Gly Arg His		
	425	430 435
Thr Phe Ser Pro		

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 436 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Asn	Thr	Thr	Leu	Pro	Ser	Trp	Lys	Asp	Arg	Thr	Gln	Asn	Gln
1				5					10					15
Phe	Gly	Lys	Leu	Gln	Ile	Gln	Val	Pro	Trp	Arg	Thr	Ile	Gln	Leu
				20					25					30
Leu	Val	Pro	His	Arg	Met	Arg	Arg	Lys	Ile	Arg	Ser	Lys	Leu	Arg
				35					40					45
Ser	Arg	Ile	Ser	Pro	Thr	Ser	Ser	Ile	Ser	Ser	Leu	Gln	Thr	Ser
				50					55					60
Phe	Ser	Pro	Val	Asp	Thr	Leu	Arg	Ser	Leu	Gln	Ser	His	Arg	Trp
				65					70					75
Thr	Leu	Tyr	Asp	Phe	Gln	Tyr	Leu	Leu	Leu	Leu	Ile	Val	Gly	Ile
				80					85					90
Phe	Ser	Leu	Ser	Val	Met	Glu	Ser	Pro	Gly	Pro	Leu	Ala	Lys	Thr
				95					100					105
Ala	Ala	Phe	Thr	Leu	Leu	Leu	Val	Ser	Leu	Leu	Leu	Pro	Ile	Thr
				110					115					120
Arg	Gln	Phe	Phe	Leu	Pro	Phe	Leu	Pro	Ile	Ala	Gly	Trp	Leu	Ile
				125					130					135
Phe	Phe	Tyr	Ala	Cys	Gln	Phe	Ile	Pro	Ser	Asp	Trp	Arg	Pro	Ala
				140					145					150
Ile	Trp	Val	Arg	Val	Leu	Pro	Ala	Leu	Glu	Asn	Ile	Leu	Tyr	Gly
				155					160					165
Ala	Asn	Ile	Ser	Asn	Ile	Leu	Ser	Ala	His	Gln	Asn	Val	Val	Leu

170	175	180
Asp Val Leu Ala Trp Leu Pro Tyr Gly	Ile Cys His Tyr Gly Ala	
185	190	195
Pro Phe Val Cys Ser Ala Ile Met Phe	Ile Phe Gly Pro Pro Gly	
200	205	210
Thr Val Pro Leu Phe Ala Arg Thr Phe	Gly Tyr Ile Ser Met Ala	
215	220	225
Ala Val Thr Ile Gln Leu Phe Phe Pro	Cys Ser Pro Pro Trp Tyr	
230	235	240
Glu Asn Leu Tyr Gly Leu Ala Pro Ala	Asp Tyr Ser Met Pro Gly	
245	250	255
Asn Pro Ala Gly Leu Ala Arg Ile Asp	Glu Leu Phe Gly Ile Asp	
260	265	270
Leu Tyr Thr Ser Gly Phe Arg Gln Ser	Pro Val Val Phe Gly Ala	
275	280	285
Phe Pro Ser Leu His Ala Ala Asp Ser	Thr Leu Ala Ala Leu Phe	
290	295	300
Met Ser Gln Val Phe Pro Arg Leu Lys	Pro Leu Phe Val Ile Tyr	
305	310	315
Thr Leu Trp Met Trp Trp Ala Thr Met	Tyr Leu Ser His His Tyr	
320	325	330
Ala Val Asp Leu Val Gly Gly Gly Leu	Leu Ala Thr Val Ala Phe	
335	340	345

Tyr Phe Ala Lys Thr Arg Phe Met Pro Arg Val Gln Asn Asp Lys			
	350	355	360
Met Phe Arg Trp Asp Tyr Asp Tyr Val Glu Tyr Gly Asp Ser Ala			
	365	370	375
Leu Asp Tyr Gly Tyr Gly Pro Ala Ser Phe Glu Gly Glu Phe Asn			
	380	385	390
Leu Asp Ser Asp Glu Trp Thr Val Gly Ser Ser Ser Ser Ile Ser			
	395	400	405
Ser Gly Ser Leu Ser Pro Val Asp Asp His Tyr Ser Trp Glu Gly			
	410	415	420
Glu Thr Leu Ala Ser Pro Ala Thr Asp Ile Glu Ser Gly Arg His			
	425	430	435
Phe			

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5					10					15
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20					25					30
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35					40					45
Val	His	Tyr	Ile	Phe	Leu	Gly	Ser	Ile	Met	Leu	Phe	Val	Phe	Ile
				50					55					60
Thr	Asn	Pro	Ala	Pro	Trp	Ile	Phe	Lys	Ile	Leu	Phe	Tyr	Cys	Phe
				65					70					75
Leu	Gly	Thr	Leu	Phe	Ile	Ile	Pro	Ala	Thr	Ser	Gln	Phe	Phe	Phe
				80					85					90
Asn	Ala	Leu	Pro	Ile	Leu	Thr	Trp	Val	Ala	Leu	Tyr	Phe	Thr	Ser
				95					100					105
Ser	Tyr	Phe	Pro	Asp	Asp	Arg	Arg	Pro	Pro	Ile	Thr	Val	Lys	Val
				110					115					120
Leu	Pro	Ala	Val	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asp
				125					130					135
Ile	Leu	Ala	Thr	Ser	Thr	Asn	Ser	Phe	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Leu	Pro	Tyr	Gly	Leu	Phe	His	Phe	Gly	Ala	Pro	Phe	Val	Val	Ala
				155					160					165
Ala	Ile	Leu	Phe	Val	Phe	Gly	Pro	Pro	Thr	Val	Leu	Gln	Gly	Tyr

170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu	Phe Gly Val Ile Met Gln	
185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp	Tyr Lys Ile Leu Tyr Gly	
200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His	Gly Ser Pro Gly Gly Leu	
215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile	Asn Met Tyr Thr Thr Ala	
230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly	Ala Phe Pro Ser Leu His	
245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu	Phe Phe Cys Tyr Cys Phe	
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala	Tyr Val Cys Trp Leu Trp	
275	280	285
Trp Ser Thr Met Tyr Leu Thr His His	Tyr Phe Val Asp Leu Met	
290	295	300
Ala Gly Ser Val Leu Ser Tyr Val Ile	Phe Gln Tyr Thr Lys Tyr	
305	310	315
Thr His Leu Pro Ile Val Asp Thr Ser	Leu Phe Cys Arg Trp Ser	
320	325	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile	Ser Lys Ser Asp Pro Leu	
335	340	345

Thr	Phe	Arg	Leu	Leu	Arg	Asn	Thr	Lys	Trp	Ser	Trp	Thr	His	Leu
				35					40					45
Gln	Tyr	Val	Phe	Leu	Ala	Gly	Asn	Leu	Ile	Phe	Ala	Cys	Ile	Val
				50					55					60
Ile	Glu	Ser	Pro	Gly	Phe	Trp	Gly	Lys	Phe	Gly	Ile	Ala	Cys	Leu
				65					70					75
Leu	Ala	Ile	Ala	Leu	Thr	Val	Pro	Leu	Thr	Arg	Gln	Ile	Phe	Phe
				80					85					90
Pro	Ala	Ile	Val	Ile	Ile	Thr	Trp	Ala	Ile	Leu	Phe	Tyr	Ser	Cys
				95					100					105
Arg	Phe	Ile	Pro	Glu	Arg	Trp	Arg	Pro	Pro	Ile	Trp	Val	Arg	Val
				110					115					120
Leu	Pro	Thr	Leu	Glu	Asn	Ile	Leu	Tyr	Gly	Ser	Asn	Leu	Ser	Ser
				125					130					135
Leu	Leu	Ser	Lys	Thr	Thr	His	Ser	Ile	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Val	Pro	Tyr	Gly	Val	Met	His	Tyr	Ser	Ala	Pro	Phe	Ile	Ile	Ser
				155					160					165
Phe	Ile	Leu	Phe	Ile	Phe	Ala	Pro	Pro	Gly	Thr	Leu	Pro	Val	Trp
				170					175					180
Ala	Arg	Thr	Phe	Gly	Tyr	Met	Asn	Leu	Phe	Gly	Val	Leu	Ile	Gln
				185					190					195
Met	Ala	Phe	Pro	Cys	Ser	Pro	Pro	Trp	Tyr	Glu	Asn	Met	Tyr	Gly

200	205	210
Leu Glu Pro Ala Thr Tyr Ala Val Arg Gly Ser Pro Gly Gly Leu		
215	220	225
Ala Arg Ile Asp Ala Leu Phe Gly Thr Ser Ile Tyr Thr Asp Gly		
230	235	240
Phe Ser Asn Ser Pro Val Val Phe Gly Ala Phe Pro Ser Leu His		
245	250	255
Ala Gly Trp Ala Met Leu Glu Ala Leu Phe Leu Ser His Val Phe		
260	265	270
Pro Arg Tyr Arg Phe Cys Phe Tyr Gly Tyr Val Leu Trp Leu Cys		
275	280	285
Trp Cys Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Val		
290	295	300
Gly Gly Met Cys Leu Ala Ile Ile Cys Phe Val Phe Ala Gln Lys		
305	310	315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu		
320	325	330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser		
335	340	345
Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp		
350	355	360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu		
365	370	375

Asn	Asn	Met	Glu	Leu	Ala	Asn	Thr	Asp	His	Glu	Trp	Ser	Val	Gly
				380					385					390
Ser	Ser	Ser	Pro	Glu	Pro	Leu	Pro	Ser	Pro	Ala	Ala	Asp	Leu	Ile
				395					400					405
Asp	Arg	Pro	Ala	Ser	Thr	Thr	Ser	Ser	Ile	Phe	Asp	Ala	Ser	His
				410					415					420
Leu	Pro													

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Ala	Ser	Ser	Ile	Leu	Arg	Ser	Lys	Ile	Ile	Gln	Lys	Pro	Tyr
1				5				10						15
Gln	Leu	Phe	His	Tyr	Tyr	Phe	Leu	Ser	Glu	Lys	Ala	Pro	Gly	Ser
				20				25						30
Thr	Val	Ser	Asp	Leu	Asn	Phe	Asp	Thr	Asn	Ile	Gln	Thr	Ser	Leu
				35				40						45

Arg	Lys	Leu	Lys	His	His	His	Trp	Thr	Val	Gly	Glu	Ile	Phe	His
				50					55					60
Tyr	Gly	Phe	Leu	Val	Ser	Ile	Leu	Phe	Phe	Val	Phe	Val	Val	Phe
				65					70					75
Pro	Ala	Ser	Phe	Phe	Ile	Lys	Leu	Pro	Ile	Ile	Leu	Ala	Phe	Ala
				80					85					90
Thr	Cys	Phe	Leu	Ile	Pro	Leu	Thr	Ser	Gln	Phe	Phe	Leu	Pro	Ala
				95					100					105
Leu	Pro	Val	Phe	Thr	Trp	Leu	Ala	Leu	Tyr	Phe	Thr	Cys	Ala	Lys
				110					115					120
Ile	Pro	Gln	Glu	Trp	Lys	Pro	Ala	Ile	Thr	Val	Lys	Val	Leu	Pro
				125					130					135
Ala	Met	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asn	Val	Leu
				140					145					150
Ala	Thr	Ile	Thr	Thr	Gly	Val	Leu	Asp	Ile	Leu	Ala	Trp	Leu	Pro
				155					160					165
Tyr	Gly	Ile	Ile	His	Phe	Ser	Phe	Pro	Phe	Val	Leu	Ala	Ala	Ile
				170					175					180
Ile	Phe	Leu	Phe	Gly	Pro	Pro	Thr	Ala	Leu	Arg	Ser	Phe	Gly	Phe
				185					190					195
Ala	Phe	Gly	Tyr	Met	Asn	Leu	Leu	Gly	Val	Leu	Ile	Gln	Met	Ala
				200					205					210
Phe	Pro	Ala	Ala	Pro	Pro	Trp	Tyr	Lys	Asn	Leu	His	Gly	Leu	Glu

215	220	225
Pro Ala Asn Tyr Ser Met His Gly Ser	Pro Gly Gly Leu Gly Arg	
230	235	240
Ile Asp Lys Leu Leu Gly Val Asp Met Tyr Thr Thr Gly Phe Ser		
245	250	255
Asn Ser Ser Ile Ile Phe Gly Ala Phe Pro Ser Leu His Ser Gly		
260	265	270
Cys Cys Ile Met Glu Val Leu Phe Leu Cys Trp Leu Phe Pro Arg		
275	280	285
Phe Lys Phe Val Trp Val Thr Tyr Ala Ser Trp Leu Trp Trp Ser		
290	295	300
Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Ile Gly Gly		
305	310	315
Ala Met Leu Ser Leu Thr Val Phe Glu Phe Thr Lys Tyr Lys Tyr		
320	325	330
Leu Pro Lys Asn Lys Glu Gly Leu Phe Cys Arg Trp Ser Tyr Thr		
335	340	345
Glu Ile Glu Lys Ile Asp Ile Gln Glu Ile Asp Pro Leu Ser Tyr		
350	355	360
Asn Tyr Ile Pro Val Asn Ser Asn Asp Asn Glu Ser Arg Leu Tyr		
365	370	375
Thr Arg Val Tyr Gln Glu Ser Gln Val Ser Pro Pro Gln Arg Ala		
380	385	390

Glu Thr Pro Glu Ala Phe Glu Met Ser Asn Phe Ser Arg Ser Arg		
	395	400 405
Gln Ser Ser Lys Thr Gln Val Pro Leu Ser Asn Leu Thr Asn Asn		
	410	415 420
Asp Gln Val Ser Gly Ile Asn Glu Glu Asp Glu Glu Glu Glu Gly		
	425	430 435
Asp Glu Ile Ser Ser Ser Thr Pro Ser Val Phe Glu Asp Glu Pro		
	440	445 450
Gln Gly Ser Thr Tyr Ala Ala Ser Ser Ala Thr Ser Val Asp Asp		
	455	460 465
Leu Asp Ser Lys Arg Asn		
	470	

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 3 is Val or Ile"

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 7 is Leu or Val"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Leu Asp Xaa Leu Ala Trp Xaa Pro Tyr Gly

1

5

10

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Phe Gly Ala Phe Pro Ser Leu His

1

5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 5 is Ser or Thr"

(ix) FEATURE:

(A) NAME/KEY:

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION: /note= "Xaa at position 9 is Ala or Phe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Thr Met Tyr Leu Xaa His His Tyr Xaa Val Asp Leu

1

5

10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2935 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
ATTTTCTTCC CCATAACAAC TCTTCTCGCC CTCCTCCGG CTCCGTGGCC AAATTGTTTT 60
ATGCAGCGCC TCCTAGCGAT TTAACCTCGT TCTCGTTGCC CTTGCCTGTC CGCCTTGCGT 120
CAGTACGACC CTTGCAACGT GACCTTCCCC AGAGTATCCT CGTTTGGCCG CTGGAGACCG 180
GAGCTTGAC C CTCATAAAC TAGCTCTTCG AAATCAATTC TCCGTTCTCC AGAGATTATC 240
GGATCGAATC TCTCCGCTGT CGACACCTTT CGTCTCTCGG TGATCCTCGC CCTTGGAGTC 300
TCGTCACGTT GACGCCTTGA ACCCCTGGCC GCCAACTCCA CATAGGAGAC CACACTTCAT 360
TCTTCCCCCG CCATAATTGC AGCACCTCC GTCTCCCTTC GAGCTCCTCC TGGATCATCA 420
AGTCCGAAAG GATTAGACTC GTCGCAGCGA TGAATACCAC CCTTCCATCC TGGGAAGGATC 480
GGACGCAAAA CCAGTTCGGC AAGCTCCAGA TCCAAGTCCC ATGGCGCACC ATACAACTTC 540
TCGTGCCGCA CCGTATGCGA CGGAAGATTC GGTCCAAGCT GCGCAGTCGG ATCTCGCCTA 600
CCTCATCGAT ATCCTCGTTG CAGACGTCAT TCTCACCTGT CGATACACTC AGGTCGCTGC 660
AAAGTCATAG ATGGACGCTC TATGACTTTC AGTATCTTTT GCTGCTGATT GTCGGCATAT 720
TCTCGCTGAG CGTTATGGAA TCACCTGGAC CATTGGCAAA GACCGCCGCG TTTACGCTAC 780
TTCTCGTCTC TCTCCTTCTC CCGATTACGC GCCAGTTCTT CTTGCCATTC CTCCCGATTG 840
CAGGATGGCT TATATTTTTT TACGCTTGCC AGTTCATCCC GAGCGACTGG CGCCCTGCAA 900
```

TCTGGGTTCG CGTGCTGCCG GCTCTGGAAA ACATTCTCTA CGGTGCTAAT ATCAGTAACA 960
TCCTTTCCGC TCACCAAAAT GTGGTGCTTG ACGTTTTGGC GTGGCTTCCC TACGGAATCT 1020
GCCATTATGG CGCGCCATTT GTGTGCTCAG CGATCATGTT CATCTTTGGT CCTCCCGGCA 1080
CCGTCCCCCT TTTCGCTCGA ACTTTTGGAT ACATCAGCAT GGCTGCAGTC ACCATTCAGC 1140
TGTTTTTCCC CTGCTCTCCT CCGTGGTACG AAAATCTGTA TGGTTTGGCT CCGGCTGATT 1200
ACTCCATGCC GGGTAATCCT GCGGGCCTTG CTCGCATCGA TGAGCTTTTTT GGGATAGACT 1260
TGTACACATC GGGCTTCAGA CAATCTCCCG TCGTGTTTGG CGCATTTCTT TCCCTACATG 1320
CCGCTGATTC GACACTTGCA GCTCTATTTA TGAGCCAAGT GTTCCCACGG TTGAAGCCCT 1380
TGTTTGTCAT CTATACTCTC TGGATGTGGT GGGCTACAAT GTATCTTTTCG CACCACTACG 1440
CTGTTGATCT GGTCGGTGGT GGCCTCTTGG CAACTGTCGC GTTCTACTTT GCTAAAACGC 1500
GGTTCATGCC TCGCGTCCAG AATGATAAGA TGTTCCGCTG GGACTIONGAT TATGTTGAGT 1560
ACGGCGATTC CGCACTCGAC TATGGGTACG GTCCAGCCAG CTCGAAGGC GAATTCAACC 1620
TTGATAGCGA TGAGTGGACC GTTGGTTCTT CGTCATCCAT TTCGTCCGGC TCCCTCAGTC 1680
CAGTTGACGA CCACTACTCT TGGGAGGGCG AGACTCTTGC CTCTCCTGCC ACCGACATCG 1740
AGTCTGGAAG GCATTTCTGA TCCTGCTCAA TGAGCCTTGA TACGTACTAC ACTGTGTACG 1800
TGCTACTGCA TTGACTAATG AGACGGCGTT TTCAAACAAA TTTTAACGAC ATGCTTGGTT 1860
ATCGCATTGA GCTGATTTTCG ACACATATAT ATGTTTAATA CGTTTTGGGG ACACTCCAGG 1920
GATTCATGAC GGTTGCTTCA TATCCCGACC TGGGGATGGA TTGACCTGGT TGTGCCCAAT 1980
TTTCTTCTGC CTAACGTTTT GATTATACAT GCATTTTTTCA CGAAACCAGC CGGCCCAGCA 2040
TGATCGTGAC CTCAATTTGA GCTCGAATCT TCCTGGGGCC TCCAGCGATA ATTCTTAATG 2100
CTCGTTCCGA GGGTGCCACA TCGGACATTC GCTTGTACAA CTTTTGCAGA ACGAACATTT 2160
TCACCGATTC CAACTTGAGT CATTCGCTTA CTACTTTCAA CTGGTCGAGA AACTTCGCTT 2220
CTTTTCAGCT CGGCTAGGTG CATAAATATT TTACATTCGT GTCGATCGCT CACATTTTAC 2280

GGCGCCTGGA AACTTGGGGG TTTCGATTTC ATTGGAAAGG ATAAACAACA TGGGCTGGGC 2340
GCCTTTTACA CGCACTACAT TCGCTTAGAA AAGTTCTGAT GCTTTTAATG ATTCTTGCAT 2400
TGGCATATAG AAAGGGGTCC TCCAGACTCG CTACTGTGGT CCTCTCTCAA CCCCACTC 2460
GCTTGCTTTT AACAGTGGAC ACCCCGTGGA GCTACGTCTC CATCAAATAT TTGGCATCAA 2520
CCGGAATCGA TGCCAGGAGG ACTGAGCTTA CTCACGGTGA CCGTCGGGTA AAAGGCGTTC 2580
ATACAGAACA TCTCCTCTAT CCTCCTGTCC CATCTCGATT CTCTGGCTGC TGGACGCAAC 2640
ACACCTCGCT GCACGTTTTT GACTTCCTAA TACGACCTAA TATCATTCTC GGTTTTCTTT 2700
GCTCTGGCTC GCGGCCGCCA TTTATATGGC GTGTCGGCTC GAGTCTGAAG TGAACTTTTT 2760
TCTCCTTTCT GGCTCCACA ACTTTCCGAT CCCTAGCAGC TTCCCGTGCA CAGCGAGGTG 2820
TTGTTGGATG ATTGTTCCAT AGCATTATCA TTATTCCTAA TCCGGTAGCG TTATGATTTA 2880
TGAAGAACAG TGATGTACAT TATTATGCGG TGATCAAAAA AAAAAAAAAA AAAAA 2935

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2856 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TAGTATACAA GTTGCCTTCA TAAATCATAA CACTACCGGA TTATAACTAA TAATACTGCT 60

GCTATGGATA	ATTTGCCCCA	ACCACACACA	CACACACACT	ATCCAAGGGA	AACGAAAAGA	120
AGAGGATAGT	GTCCATGAGC	GGAATTTGTG	AGGTGAGGAG	GATGTGATAG	GTAGGTGTGA	180
GAGCGTCAGA	CCATAGCAAG	CAAGTCAAGT	CACCGTCCAA	CAAGTCGAGA	ATCTGAGGTT	240
AGTCGAAGCC	GGGTCAAATC	GGGACGCGTG	AACATCCTTC	AACAAGCCCC	GTCAGCGTCA	300
GA CTCTCCTC	CCAACCTTGC	ACGTCCGGCG	AATGTCAGAA	AGCAACACAG	TAGCCATGTG	360
TATAGCCCCC	GGGTAATATG	TGGACGTAAT	CCACTTTGAG	CGAATGGATA	AACCATGCGA	420
GCAATCAAGA	AGAAGGTCCA	TCCTTCCCCT	CTTCTATAT	GCTAATGCAG	AGTTCGTGAT	480
GCATTGCAAA	GCAGAAATGT	AAAACCGTCT	ATGCACAGCG	AGTGAAGTGC	AAGAAGAAAA	540
GGCGCCCCAT	TCCGTGATTT	TTTCCAATAT	AAATGCTTCC	AGTCCCAAGC	GCCAAGAGTA	600
TGAGCGATCG	ACACGAATGG	TAAAAATAAA	TATATAGCAT	AGTCGAAACG	AATAATAAGG	660
ATGGGTTCGA	CCAGGTCAAG	GTAATAGCGG	ACGATAGAAG	TGAGAATTGA	AGTAGTGTCT	720
GGGTGGACAC	GAGACAAGCG	AGTGTTACCG	ACGGGGGTTA	TCGGATGGTC	TTCCAGGACT	780
AAACACCAGG	ATGCGTAGCC	CCATTGTGTC	CATTAACATT	ATGATGGAGC	AACGGCGTGT	840
AGCGATGAGC	GATAACCCAT	CCCACGTTCG	AGCCCAACCC	GGTTTCGCAA	TGTCAAGCAA	900
TGTACAGTGA	AGTGCGGCGG	AAAGAGAAAC	AGTATCCCAG	GCGGATTCGG	ACCGAGGTAC	960
ATTAAGGTCG	AACCGATGAT	CTGGGACACC	GAAGAAAGGG	TATTAAAGAT	ATATGTGTCTG	1020
AAATGTGCTC	GATGCAATAA	AGCCAGAATG	TCGTGAAATG	TTCTAACGCC	GTCTCGTAAA	1080
TAGATGCAGT	GGAGGAGGAG	CGAGATCGCT	ATATGCAGGT	ATCGAGTTTG	GTTTGTGGCT	1140
ACTCAAGGGC	TGAAAGTATG	CCTGCCGGAC	TCAATATCAG	TATGTGGGGA	GGTCAGTGCC	1200
TCGGTTTCCC	ATGAGTAATG	ATCGTCAACG	GGACTCAAGG	AGCCTGAGGA	GACGGAGGAT	1260
GAAGAACCAA	CAGTCCATTC	ATCGCTGTCTG	AGATTGAAGT	CTCCATCATA	GCCAGCTGCA	1320
CCATACCCAT	ACTCCAGGGC	AGACTCGCCG	AATTCCACAT	AGTCGTAGTC	CCAACGGAAG	1380
GTCTTGTCGA	GCTGGACACG	GGGAAGGAAT	CGGGTCTTGG	CGAAGTAGAA	AGCAATGGCG	1440

GCCAGGAGAC CACCCGCAAC CAAATCGACC GCATAGTGAT GTGAGAGGTA CATTGTTGCC 1500
CACCACATCC ATAGAGTATA GGTCACGAAG ACGGGCTTCA TGCGGGGGAA AACATGACTC 1560
ATGAAAAGTG CGGCCAGGGT TGAGTCGGCA GCATGCAGCG ACGGAAAAGC GCCGAACACA 1620
ACAGGCGACT GATGGAAAAC AGACGTGTAA AGGTTCGATGC CGAAAAGCTT GTCAATGCGG 1680
GCAAGCCCTG CGGGATCACC TTGGATGGAG TAGTCTGCCG GAGCTAGACC ATAGCGATTC 1740
TCATACCAAG GTGGAGAGCA AGGGAAAAAC AGCTGAATAG TAACCGCAGT CATACTGATA 1800
TAGCCGAAAG TGC GCGCGAA AAGGGGAACA GTGCCGGGCG GACCGAAGAT GAACATGATC 1860
AACGAGCACA CAAACGGAGC GCCATAGTGG CAGATACCGT AGGGTAGCCA CGCCAGCACG 1920
TCAAGCACAA CGTTCTGGTG AGCGGATAGG ATGTTGCTGA TGTTTGCGCC GTAGAGAATA 1980
TTCTCCAGTG CAGGCAAGAC ACGAACCCAA ATCGCAGGGC GCCAATCGCT TGGGATGAAC 2040
TGGCAGGCGT AGAAAAACAG AAGCCATCCG GCAATCGGCA GAAACGGGAG GAAGAACTGG 2100
CGGGTCATAG GGATCAGGAG AGAGAATAGG AGCATGGAGA AAATGGCCGT TTTGCCCAA 2160
GGCCCGGGCG ACTCGATAAC GGTCAAAGAG AAGATGCCCA CGATCAACAG AAGCAGATAT 2220
TGGAAGTCGT AAACCGTCCA TCGGTGGCTT TGGAGCGATC GTAGTGTGTC TGCAGGCGAT 2280
AACGACGTCT GTAAAGAGGC TATTGACGAG GTAGGAGACG CTCTACTGCG CAATTTGGAC 2340
CTTAACTTCC GCCGCATGCG ATGCGGGACG AGCAGTTGGA TGGACCGCCA TGGAACCTGG 2400
ATCTGAAGCT TTCCAAACTG GTTCTGCGTG CGGTCCTTCC ACGTGGAAG TGTTTGATTC 2460
ATGGTTGCGG CGAGTCTTGC GGGCAAGACT CAGCTATAAT TCCTCCAGAT CGAACTCCGT 2520
CTGCTTTCTT TCCCCCGAT TGTACAAGAG TTGGAAGGCG ACGAATAAAT CGGTCGCGGA 2580
AGCGGCAGAA AAGGGGTCTT GAATAACAAG CCGTTAGGCA ACGAGCAGGG TGCTGACGGC 2640
GTCAACCCGA GCGGGGAAAT TGAGATCGGT TCGAAGAAGA CGAAACGAAG CGCACGCTCC 2700
GGTCTCCAGC GGCCAAACGA GGATGTCTTT GAGCCGTCAC GTTGCAGGGG GTTTCACTGA 2760
TGCAAGGCGA AGTGCAAGGG CGCAGCAGTC GAGAGGAAAA CCTGGAACGA GGAGGTCGTG 2820

AGGGAGGCAG ATGGCCACGG AGCCGGAGTA TAAACC

2856

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2856 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: mRNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
UAGUAUACAA GUUGCCUUCA UAAAUCAUAA CACUACCGGA UUUAACUAA UAAUACUGCU   60
GCUAUGGAUA AUUUGCCCCA ACCACACACA CACACACACU AUCCAAGGGA AACGAAAAGA  120
AGAGGAUAGU GUCCAUGAGC GGAAUUUGUG AGGUGAGGAG GAUGUGAUAG GUAGGUGUGA  180
GAGCGUCAGA CCAUAGCAAG CAAGUCAAGU CACCGUCCAA CAAGUCGAGA AUCUGAGGUU  240
AGUCGAAGCC GGGUCAAAUC GGGACGCGUG AACAUCCUUC AACAAGCCCC GUCAGCGUCA  300
GACUCUCCUC CCAACCUUGC ACGUCCGGCG AAUGUCAGAA AGCAACACAG UAGCCAUGUG  360
UAUAGCCCCC GGGUAAUAUG UGGACGUAAU CCACUUUGAG CGAAUGGAUA AACCAUGCGA  420
GCAAUCAAGA AGAAGGUCCA UCCUUCCCCU CUUUCUAUUAU GCUAAUGCAG AGUUCGUGAU  480
GCAUUGCAAA GCAGAAAUGU AAAACCGUCU AUGCACAGCG AGUGAAGUGC AAGAAGAAAA  540
GGCGCCCCAU UCCGUGAUUU UUUCCAAUAU AAAUGCUUCC AGUCCCAAGC GCCAAGAGUA  600
UGAGCGAUCG ACACGAAUGG UAAAAAUAAA UAUAUAGCAU AGUCGAAACG AAUAAUAAGG  660
```

AUGGGUUCGA	CCAGGUCAAG	GUAAUAGCGG	ACGAUAGAAG	UGAGAAUUGA	AGUAGUGUCU	720
GGGUGGACAC	GAGACAAGCG	AGUGUUACCG	ACGGGGGUUA	UCGGAUGGUC	UUCCAGGACU	780
AAACACCAGG	AUGCGUAGCC	CCAUUGUGUC	CAUUAACAUU	AUGAUGGAGC	AACGGCGUGU	840
AGCGAUGAGC	GAUAACCCAU	CCCACGUUCG	AGCCCAACCC	GGUUUCGCAA	UGUCAAGCAA	900
UGUACAGUGA	AGUGCGGCGG	AAAGAGAAAC	AGUAUCCCAG	GCGGAUUCGG	ACCGAGGUAC	960
AUUAAGGUCG	AACCGAUGAU	CUGGGACACC	GAAGAAAGGG	UAUUAAGAU	AUAUGUGUCG	1020
AAAUGUGCUC	GAUGCAAUAA	AGCCAGAAUG	UCGUGAAAUG	UUCUAACGCC	GUCUCGUAAA	1080
UAGAUGCAGU	GGAGGAGGAG	CGAGAU CGCU	AUAUGCAGGU	AUCGAGUUUG	GUUUGUGGCU	1140
ACUCAAGGGC	UGAAAGUAUG	CCUGCCGGAC	UCAUAUUCAG	UAUGUGGGGA	GGUCAGUGCC	1200
UCGGUUUCCC	AUGAGUAAUG	AUCGUCAACG	GGACUCAAGG	AGCCUGAGGA	GACGGAGGAU	1260
GAAGAACCAA	CAGUCCAUUC	AUCGCUGUCG	AGAUUGAAGU	CUCCAUCAUA	GCCAGCUGCA	1320
CCAUACCCAU	ACUCCAGGGC	AGACUCGCCG	AAUUCCACAU	AGUCGUAGUC	CCAACGGAAG	1380
GUCUUGUCGA	GCUGGACACG	GGGAAGGAAU	CGGGUCUUGG	CGAAGUAGAA	AGCAAUGGCG	1440
GCCAGGAGAC	CACCCGCAAC	CAAUUCGACC	GCAUAGUGAU	GUGAGAGGUA	CAUUGUUGCC	1500
CACCACAUCC	AUAGAGUAUA	GGUCACGAAG	ACGGGCUUCA	UGC GGGGGA	AACAUGACUC	1560
AUGAAAAGUG	CGGCCAGGGU	UGAGUCGGCA	GCAUGCAGCG	ACGGAAAAGC	GCCGAACACA	1620
ACAGGCGACU	GAUGGAAAAC	AGACGUGUAA	AGGUCGAUGC	CGAAAAGCUU	GUCAAUGCGG	1680
GCAAGCCCUG	CGGGAUCACC	UUGGAUGGAG	UAGUCUGCCG	GAGCUAGACC	AUAGCGAUUC	1740
UCAUACCAAG	GUGGAGAGCA	AGGGAAAAAC	AGCUGAAUAG	UAACCGCAGU	CAUACUGAUA	1800
UAGCCGAAAG	UGC GCGCGAA	AAGGGGAACA	GUGCCGGGCG	GACCGAAGAU	GAACAUGAUC	1860
AACGAGCACA	CAAACGGAGC	GCCAUAGUGG	CAGAUACCGU	AGGGUAGCCA	CGCCAGCACG	1920
UCAAGCACAA	CGUUCUGGUG	AGCGGAUAGG	AUGUUGCUGA	UGUUUGCGCC	GUAGAGAAUA	1980
UUCUCCAGUG	CAGGCAAGAC	ACGAACCCAA	AUCGCAGGGC	GCCAAUCGCU	UGGGAUGAAC	2040

```

UGGCAGGCGU AGAAAAACAG AAGCCAUCCG GCAAUCGGCA GAAACGGGAG GAAGAACUGG 2100
CGGGUCAUAG GGAUCAGGAG AGAGAAUAGG AGCAUGGAGA AAAUGGCCGU UUUGCCCAA 2160
GGCCCGGGCG ACUCGAUAAC GGUCAAAGAG AAGAUGCCCA CGAUCAACAG AAGCAGAUAU 2220
UGGAAGUCGU AAACCGUCCA UCGGUGGCUU UGGAGCGAUC GUAGUGUGUC UGCAGGCGAU 2280
AACGACGUCU GUAAAGAGGC UAUUGACGAG GUAGGAGACG CUCUACUGCG CAAUUUGGAC 2340
CUUAACUUC GCGCAUGCG AUGCGGGACG AGCAGUUGGA UGGACCGCCA UGGAACCUGG 2400
AUCUGAAGCU UUCCAAACUG GUUCUGCGUG CGGUCCUUC ACGUGGGAAG UGUUUGAUUC 2460
AUGGUUGCGG CGAGUCUUGC GGGCAAGACU CAGCUAUAAU UCCUCCAGAU CGAACUCCGU 2520
CUGCUUUCU UCCCCCGAU UGUACAAGAG UUGGAAGGCG ACGAAUAAU CGGUCGCGGA 2580
AGCGGCAGAA AAGGGGUCUU GAAUAACAAG CCGUUAGGCA ACGAGCAGGG UGCUGACGGC 2640
GUCAACCCGA GCGGGGAAAU UGAGAUCGGU UCGAAGAAGA CGAAACGAAG CGCACGCUCC 2700
GGUCUCCAGC GGCCAAACGA GGAUGUCUUU GAGCCGUCAC GUUGCAGGGG GUUUCACUGA 2760
UGCAAGGCGA AGUGCAAGGG CGCAGCAGUC GAGAGGAAAA CCUGGAACGA GGAGGUCGUG 2820
AGGGAGGCAG AUGGCCACGG AGCCGGAGUA UAAACC 2856

```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGCTTTTTT	GCCTCTGCAA	AAGTTCCTTT	CTCGAATTGG	TTTTTTGAGG	AAAAGCAAGT	60
TAATAAACTA	ATTATATTAT	ATATAATTAG	CAATTTTATA	AAAAAAATAA	AAAAATAGCC	120
CTGATTGCTG	GCAACTGTGA	GCTGAACATT	GGTTAATCGG	TCCATCTTTT	TTTAAATATT	180
TTACATCGCT	ACTTTTAAGT	GCTTGACACT	TGCATTTAAT	AGCTACTTTC	TTTCCTTCAT	240
AAAAATTCCT	TTTTTTTCCT	TTAGTTTTCC	GGTTAATTCC	TTACGAAATT	TTTTTCGTAC	300
GCTTCCCTTT	TTTACTCTGA	TAATTCTTTG	AAGCAATGTC	TGCTCTTTCG	ACCTTAAAAA	360
AGCGCCTTGC	TGCGTGTAAC	CGAGCATCCC	AATACAAGTT	GGAAACAAGC	TTAAACCCTA	420
TGCCTACATT	TCGTTTGCTA	CGCAATACGA	AATGGTCATG	GACACATTTG	CAATATGTGT	480
TTCTAGCAGG	TAATTTGATT	TTTGCTTGTA	TTGTCATTGA	ATCTCCTGGA	TTCTGGGGGA	540
AATTTGGCAT	TGCCTGTCTT	TTGGCCATTG	CGTTGACCGT	TCCTTTAACA	CGCCAAATTT	600
TTTTTCCTGC	CATTGTTATC	ATCACCTGGG	CAATTTTATT	TTACTCTTGT	AGGTTTATTC	660
CAGAACGCTG	GCGTCCACCC	ATATGGGTTC	GTGTTTTACC	CACACTTGAA	AATATTCTTT	720
ATGGCTCTAA	TCTTTCTAGT	CTTCTCTCGA	AAACCACGCA	TAGCATCCTT	GATATTTTGG	780
CCTGGGTTCC	ATATGGAGTC	ATGCATTATT	CGGCTCCTTT	TATCATTTCA	TTTATTCTTT	840
TCATCTTTGC	ACCTCCTGGA	ACTCTTCCAG	TTTGGGCTCG	AACTTTTGGT	TATATGAATT	900
TATTTGGTGT	TCTTATCCAA	ATGGCTTTCC	CCTGTTCTCC	TCCTTGGTAT	GAAAATATGT	960
ATGGTTTAGA	ACCTGCCACG	TATGCAGTAC	GTGGCTCTCC	TGGTGGATTG	GCCCGTATTG	1020
ATGCTCTCTT	CGGCACTAGC	ATTTACACTG	ATTGTTTTTC	TAACCTCTCCG	GTTGTTTTTG	1080
GTGCCTTTCC	ATCTCTTCAC	GCTGGATGGG	CCATGCTGGA	AGCACTTTTC	CTTTCGCATG	1140
TGTTTCCTCG	ATACCGCTTC	TGCTTTTATG	GATATGTTCT	ATGGCTTTGC	TGGTGTACTA	1200
TGTACCTTAC	CCACCACTAC	TTTGTAGATT	TGGTCGGCGG	TATGTGTTTA	GCTATTATAT	1260
GCTTCGTTTT	TGCTCAAAAG	CTACGCCTCC	CACAGTTGCA	AACTGGTAAA	ATCCTTCGTT	1320

GGGAATACGA	GTTTGTATC	CACGGTCATG	GTCTTTCCGA	AAAAACCAGC	AACTCCTTGG	1380
CTCGTACCGG	CAGCCCATAC	TTACTTGGAA	GGGATTCTTT	TACTCAAAAC	CCTAATGCAG	1440
TAGCCTTCAT	GAGTGGTCTT	AACAATATGG	AACTTGCTAA	CACCGATCAT	GAATGGTCCG	1500
TGGGTTCATC	ATCACCTGAG	CCGTTACCTA	GTCCTGCTGC	TGATTTGATT	GATCGTCCTG	1560
CCAGTACCAC	TTCCTCCATC	TTTGATGCAA	GTCATCTTCC	TTAAATCAAC	GTGCTTTAAG	1620
AATATATTTT	CAAAAGCTAC	ATGATACATT	GACTAGAATC	GGTTTGATT	ATAGTGGTAT	1680
TGGAATGATG	TTGTTCATTG	TGTTTTTTAA	CTGTTAATCT	GACATCCATT	GAGTCATTCT	1740
TTACAATTTG	TAAAATTAAT	TTGTATCACT	AATTTTGAAG	GAAGCTATTT	TGGTATTAAT	1800
ACCGCTTTTG	GTCTCCACTT	CCTTTTCGAA	ACTCTTAACA	GCGATTAGGC	CGGGTATCTT	1860
CCAGTGTGAT	GTATAGGTAT	TTGTCGTTTT	TTTATCATTT	CCGTTAATAA	AGAACTCTTT	1920
TATCCAGCTT	CTTACACTGT	CAACTGTTGT	GAAAGGAACA	CATTTAGAAT	TTCATTTTCC	1980
TTATTTGTTG	TGATTTAAAT	CGTTTGACAT	AATTTTAAAT	TTGGTTTGAA	ATGTGTGTGA	2040
GAAGGCTTGT	TTTATTCATT	TAGTTTATTG	CTTGTTTGCA	CGAAAATCCA	GAACGGAGCA	2100
TTAATGTAAT	CCTTTTTTAT	TCTGTAAAGC	GTTTTTATAC	AAATGTTGGT	TATACGTTTC	2160
TAAAATAAGA	ATATTGTTAT	AATAATATAG	TTTTTTCTAT	CATTTGTTAC	ACACACTAAA	2220
GAGACATTAA	GGATAAGCAA	ATGTGTTAAA	ATGATAATAT	ATTTTGGAAA	CATTTATAAA	2280
GAAATTAAGC	AGCTTTGACT	AACTACATTT	TTGTTTTTTT	CCTAAGCAAA	ACTGTATAGT	2340
TATACACGCG	AGCTGTATTC	ACTTCCATTG	TAGTGAATTG	AGCTC		2385

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 422 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Ser	Ala	Leu	Ser	Thr	Leu	Lys	Lys	Arg	Leu	Ala	Ala	Cys	Asn
1				5					10					15
Arg	Ala	Ser	Gln	Tyr	Lys	Leu	Glu	Thr	Ser	Leu	Asn	Pro	Met	Pro
				20					25					30
Thr	Phe	Arg	Leu	Leu	Arg	Asn	Thr	Lys	Trp	Ser	Trp	Thr	His	Leu
				35					40					45
Gln	Tyr	Val	Phe	Leu	Ala	Gly	Asn	Leu	Ile	Phe	Ala	Cys	Ile	Val
				50					55					60
Ile	Glu	Ser	Pro	Gly	Phe	Trp	Gly	Lys	Phe	Gly	Ile	Ala	Cys	Leu
				65					70					75
Leu	Ala	Ile	Ala	Leu	Thr	Val	Pro	Leu	Thr	Arg	Gln	Ile	Phe	Phe
				80					85					90
Pro	Ala	Ile	Val	Ile	Ile	Thr	Trp	Ala	Ile	Leu	Phe	Tyr	Ser	Cys
				95					100					105
Arg	Phe	Ile	Pro	Glu	Arg	Trp	Arg	Pro	Pro	Ile	Trp	Val	Arg	Val
				110					115					120
Leu	Pro	Thr	Leu	Glu	Asn	Ile	Leu	Tyr	Gly	Ser	Asn	Leu	Ser	Ser
				125					130					135

Leu	Leu	Ser	Lys	Thr	Thr	His	Ser	Ile	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Val	Pro	Tyr	Gly	Val	Met	His	Tyr	Ser	Ala	Pro	Phe	Ile	Ile	Ser
				155					160					165
Phe	Ile	Leu	Phe	Ile	Phe	Ala	Pro	Pro	Gly	Thr	Leu	Pro	Val	Trp
				170					175					180
Ala	Arg	Thr	Phe	Gly	Tyr	Met	Asn	Leu	Phe	Gly	Val	Leu	Ile	Gln
				185					190					195
Met	Ala	Phe	Pro	Cys	Ser	Pro	Pro	Trp	Tyr	Glu	Asn	Met	Tyr	Gly
				200					205					210
Leu	Glu	Pro	Ala	Thr	Tyr	Ala	Val	Arg	Gly	Ser	Pro	Gly	Gly	Leu
				215					220					225
Ala	Arg	Ile	Asp	Ala	Leu	Phe	Gly	Thr	Ser	Ile	Tyr	Thr	Asp	Cys
				230					235					240
Phe	Ser	Asn	Ser	Pro	Val	Val	Phe	Gly	Ala	Phe	Pro	Ser	Leu	His
				245					250					255
Ala	Gly	Trp	Ala	Met	Leu	Glu	Ala	Leu	Phe	Leu	Ser	His	Val	Phe
				260					265					270
Pro	Arg	Tyr	Arg	Phe	Cys	Phe	Tyr	Gly	Tyr	Val	Leu	Trp	Leu	Cys
				275					280					285
Trp	Cys	Thr	Met	Tyr	Leu	Thr	His	His	Tyr	Phe	Val	Asp	Leu	Val
				290					295					300
Gly	Gly	Met	Cys	Leu	Ala	Ile	Ile	Cys	Phe	Val	Phe	Ala	Gln	Lys

305	310	315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu		
320	325	330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser		
335	340	345
Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp		
350	355	360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu		
365	370	375
Asn Asn Met Glu Leu Ala Asn Thr Asp His Glu Trp Ser Val Gly		
380	385	390
Ser Ser Ser Pro Glu Pro Leu Pro Ser Pro Ala Ala Asp Leu Ile		
395	400	405
Asp Arg Pro Ala Ser Thr Thr Ser Ser Ile Phe Asp Ala Ser His		
410	415	420
Leu Pro		

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
AAGCTTTTTT GCCTCTGCAA AAGTTCCTTT CTCGAATTGG TTTTTTGAGG AAAAGCAAGT   60
TAATAAACTA ATTATATTAT ATATAATTAG CAATTTTATA AAAAAAATAA AAAAATAGCC  120
CTGATTGCTG GCAACTGTGA GCTGAACATT GGTTAATCGG TCCATCTTTT TTTAAATATT  180
TTACATCGCT ACTTTTAAGT GCTTGACACT TGCATTTAAT AGCTACTTTC TTTCCCTTCAT  240
AAAAATTCCT TTTTTTTCCT TTAGTTTTCC GGTTAATTCC TTACGAAATT TTTTTCGTAC  300
GCTTCCCTTT TTTACTCTGA TAATTCTTTG AAGCAATGTC TGCTCTTTCG ACCTTAAAAA  360
AGCGCCTTGC TGCCTGTAAC CGAGCATCCC AATACAAGTT GGAAACAAGC TTAAACCCTA  420
TGCCTACATT TCGTTTGCTA CGCAATACGA AATGGTCATG GACACATTTG CAATATGTGT  480
TTCTAGCAGG TAATTTGATT TTTGCTTGTA TTGTCATTGA ATCTCCTGGA TTCTGGGGGA  540
AATTTGGCAT TGCCTGTCTT TTGGCCATTG CGTTGACCGT TCCTTTAACA CGCCAAATTT  600
TTTTTCCTGC CATTGTTATC ATCACCTGGG CAATTTTATT TTACTCTTGT AGGTTTATTC  660
CAGAACGCTG GCGTCCACCC ATATGGGTTC GTGTTTTACC CACACTTGAA AATATTCTTT  720
ATGGCTCTAA TCTTTCTAGT CTTCTCTCGA AAACCACGCA TAGCATCCTT GATATTTTGG  780
CCTGGGTTCC ATATGGAGTC ATGCATTATT CGGCTCCTTT TATCATTTCA TTTATTCTTT  840
TCATCTTTGC ACCTCCTGGA ACTCTTCCAG TTTGGGCTCG AACTTTTGGT TATATGAATT  900
TATTTGGTGT TCTTATCCAA ATGGCTTTCC CCTGTTCTCC TCCTTGGTAT GAAAATATGT  960
ATGGTTTAGA ACCTGCCACG TATGCAGTAC GTGGCTCTCC TGGTGGATTG GCCCGTATTG 1020
ATGCTCTCTT CGGCACTAGC ATTTACACTG ATGGTTTTTC TAACTCTCCG GTTGTTTTTTG 1080
GTGCCTTTCC ATCTCTTCAC GCTGGATGGG CCATGCTGGA AGCACTTTTC CTTTCGCATG 1140
TGTTTCCTCG ATACCGCTTC TGCTTTTATG GATATGTTCT ATGGCTTTGC TGGTGTACTA 1200
```

TGTACCTTAC CCACCACTAC TTTGTAGATT TGGTCGGCGG TATGTGTTTA GCTATTATAT 1260
GCTTCGTTTT TGCTCAAAAG CTACGCCTCC CACAGTTGCA AACTGGTAAA ATCCTTCGTT 1320
GGGAATACGA GTTTGTTATC CACGGTCATG GTCTTTCCGA AAAAACCAGC AACTCCTTGG 1380
CTCGTACCGG CAGCCCATAC TTACTTGGA A GGGATTCTTT TACTCAAAAC CCTAATGCAG 1440
TAGCCTTCAT GAGTGGTCTT AACAAATATGG AACTTGCTAA CACCGATCAT GAATGGTCCG 1500
TGGGTTCATC ATCACCTGAG CCGTTACCTA GTCCTGCTGC TGATTTGATT GATCGTCCTG 1560
CCAGTACCAC TTCCTCCATC TTTGATGCAA GTCATCTTCC TTAAATCAAC GTGCTTTAAG 1620
AATATATTTT CAAAAGCTAC ATGATACATT GACTAGAATC GGTTTGATT C ATAGTGGTAT 1680
TGGAATGATG TTGTTCAATTG TGTTTTTTTAA CTGTTAATCT GACATCCATT GAGTCATTCT 1740
TTACAATTTG TAAAATTAAT TTGTATCACT AATTTTGAAG GAAGCTATTT TGGTATTAAT 1800
ACCGCTTTTG GTCTCCACTT CCTTTTCGAA ACTCTTAACA GCGATTAGGC CGGGTATCTT 1860
CCAGTGTGAT GTATAGGTAT TTGTCGTTTT TTTATCATTT CCGTTAATAA AGAACTCTTT 1920
TATCCAGCTT CTTACACTGT CAACTGTTGT GAAAGGAACA CATTTAGAAT TTCATTTTCC 1980
TTATTTGTTG TGATTTAAAT CGTTTGACAT AATTTTAAAT TTGGTTTGAA ATGTGTGTGA 2040
GAAGGCTTGT TTTATTCATT TAGTTTATTG CTTGTTTGCA CGAAAATCCA GAACGGAGCA 2100
TTAATGTAAT CCTTTTTTAT TCTGTAAAGC GTTTTTTATAC AAATGTTGGT TATACGTTTC 2160
TAAAATAAGA ATATTGTTAT AATAATATAG TTTTTTCTAT CATTTGTTAC ACACACTAAA 2220
GAGACATTAA GGATAAGCAA ATGTGTTAAA ATGATAATAT ATTTTGGAAA CATTTATAAA 2280
GAAATTAAGC AGCTTTGACT AACTACATTT TTGTTTTTTT CCTAAGCAAA ACTGTATAGT 2340
TATACACGCG AGCTGTATTC ACTTCCATTG TAGTGA CTG AGCTC 2385

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 422 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ser	Ala	Leu	Ser	Thr	Leu	Lys	Lys	Arg	Leu	Ala	Ala	Cys	Asn
1				5					10					15
Arg	Ala	Ser	Gln	Tyr	Lys	Leu	Glu	Thr	Ser	Leu	Asn	Pro	Met	Pro
				20					25					30
Thr	Phe	Arg	Leu	Leu	Arg	Asn	Thr	Lys	Trp	Ser	Trp	Thr	His	Leu
				35					40					45
Gln	Tyr	Val	Phe	Leu	Ala	Gly	Asn	Leu	Ile	Phe	Ala	Cys	Ile	Val
				50					55					60
Ile	Glu	Ser	Pro	Gly	Phe	Trp	Gly	Lys	Phe	Gly	Ile	Ala	Cys	Leu
				65					70					75
Leu	Ala	Ile	Ala	Leu	Thr	Val	Pro	Leu	Thr	Arg	Gln	Ile	Phe	Phe
				80					85					90
Pro	Ala	Ile	Val	Ile	Ile	Thr	Trp	Ala	Ile	Leu	Phe	Tyr	Ser	Cys
				95					100					105
Arg	Phe	Ile	Pro	Glu	Arg	Trp	Arg	Pro	Pro	Ile	Trp	Val	Arg	Val
				110					115					120

Leu	Pro	Thr	Leu	Glu	Asn	Ile	Leu	Tyr	Gly	Ser	Asn	Leu	Ser	Ser
				125					130					135
Leu	Leu	Ser	Lys	Thr	Thr	His	Ser	Ile	Leu	Asp	Ile	Leu	Ala	Trp
				140					145					150
Val	Pro	Tyr	Gly	Val	Met	His	Tyr	Ser	Ala	Pro	Phe	Ile	Ile	Ser
				155					160					165
Phe	Ile	Leu	Phe	Ile	Phe	Ala	Pro	Pro	Gly	Thr	Leu	Pro	Val	Trp
				170					175					180
Ala	Arg	Thr	Phe	Gly	Tyr	Met	Asn	Leu	Phe	Gly	Val	Leu	Ile	Gln
				185					190					195
Met	Ala	Phe	Pro	Cys	Ser	Pro	Pro	Trp	Tyr	Glu	Asn	Met	Tyr	Gly
				200					205					210
Leu	Glu	Pro	Ala	Thr	Tyr	Ala	Val	Arg	Gly	Ser	Pro	Gly	Gly	Leu
				215					220					225
Ala	Arg	Ile	Asp	Ala	Leu	Phe	Gly	Thr	Ser	Ile	Tyr	Thr	Asp	Gly
				230					235					240
Phe	Ser	Asn	Ser	Pro	Val	Val	Phe	Gly	Ala	Phe	Pro	Ser	Leu	His
				245					250					255
Ala	Gly	Trp	Ala	Met	Leu	Glu	Ala	Leu	Phe	Leu	Ser	His	Val	Phe
				260					265					270
Pro	Arg	Tyr	Arg	Phe	Cys	Phe	Tyr	Gly	Tyr	Val	Leu	Trp	Leu	Cys
				275					280					285
Trp	Cys	Thr	Met	Tyr	Leu	Thr	His	His	Tyr	Phe	Val	Asp	Leu	Val

290	295	300
Gly Gly Met Cys Leu Ala Ile Ile Cys Phe Val Phe Ala Gln Lys		
305	310	315
Leu Arg Leu Pro Gln Leu Gln Thr Gly Lys Ile Leu Arg Trp Glu		
320	325	330
Tyr Glu Phe Val Ile His Gly His Gly Leu Ser Glu Lys Thr Ser		
335	340	345
Asn Ser Leu Ala Arg Thr Gly Ser Pro Tyr Leu Leu Gly Arg Asp		
350	355	360
Ser Phe Thr Gln Asn Pro Asn Ala Val Ala Phe Met Ser Gly Leu		
365	370	375
Asn Asn Met Glu Leu Ala Asn Thr Asp His Glu Trp Ser Val Gly		
380	385	390
Ser Ser Ser Pro Glu Pro Leu Pro Ser Pro Ala Ala Asp Leu Ile		
395	400	405
Asp Arg Pro Ala Ser Thr Thr Ser Ser Ile Phe Asp Ala Ser His		
410	415	420
Leu Pro		

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2340 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: genomic DNA
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

TTTCTTTCTG TCAAAGAATA ATAAAGTGCC CATCAGTGTT CATATTTGTT ACAAAGTGGT   60
TTTCTGATTT GGTACTACTG CAGAGGCGTA TTTTTTGCTT CAGTTACCAT AGCGTAAGAA  120
CACTAGCGAC TTTTGTTTCGT GAACCAACAG AGTAGGATTT CTACTGCTAC ATCTCTTAGG  180
TAGTTGGTTA GTCCGATCGC TCACTTTTGG TTGTTGTAA GTACTTCATA AGTTTATCCT  240
TTTCCTTTTT CACACTGAGC TACTTTGGGT ATAGCTTTTG GCCCAAGGAT CTTTGAATTT  300
TCTCCAAAAG TACTTTATTT TATATCCTAC AGGTTGCGGT TTCATATTT TAAAAAGCTT  360
TTTAATCATT CCTTTGCGTA TGGCAAACCC TTTTTCGAGA TGGTTTCTAT CAGAGAGACC  420
TCCAAACTGC CATGTAGCCG ATTTAGAAAC AAGTTTAGAT CCCCATCAA CGTTGTTGAA  480
GGTGCAAAAA TACAAACCCG CTTTAAGCGA CTGGGTGCAT TACATCTTCT TGGGATCCAT  540
CATGCTGTTT GTGTTCATTA CTAATCCCGC ACCTTGGATC TTCAAGATCC TTTTTTATTG  600
TTTCTTGGGC ACTTTATTCA TCATTCCAGC TACGTCACAG TTTTCTTCA ATGCCTTGCC  660
CATCCTAACA TGGGTGGCGC TGTATTTTAC TTCATCGTAC TTCCAGATG ACCGCAGGCC  720
TCCTATTACT GTCAAAGTGT TACCAGCGGT GGAAACAATT TTATACGGCG ACAATTTAAG  780
TGATATTCTT GCAACATCGA CGAATTCCTT TTTGGACATT TTAGCATGGT TACCGTACGG  840
ACTATTTTCA TATGGGGCCC CATTTGTCGT TGCTGCCATC TTATTCGTAT TTGGTCCACC  900
AACTGTTTTG CAAGGTTATG CTTTTGCATT TGGTTATATG AACCTGTTTG GTGTTATCAT  960
GCAAAATGTC TTTCCAGCCG CTCCCCCATG GTATAAAATT CTCTATGGAT TGCAATCAGC 1020
CAACTATGAT ATGCATGGCT CGCCTGGTGG ATTAGCTAGA ATTGATAAGC TACTCGGTAT 1080

```

TAATATGTAT ACTACAGCTT TTTCAAATTC CTCCGTCATT TTCGGTGCTT TTCCTTCACT 1140
 GCATTCCGGG TGTGCTACTA TGGAAGCCCT GTTTTTCTGT TATTGTTTTTC CAAAATTGAA 1200
 GCCCTTGTTT ATTGCTTATG TTTGCTGGTT ATGGTGGTCA ACTATGTATC TGACACACCA 1260
 TTATTTTGTA GACCTTATGG CAGGTTCTGT GCTGTCATAC GTTATTTTCC AGTACACAAA 1320
 GTACACACAT TTACCAATTG TAGATACATC TCTTTTTTGC AGATGGTCAT ACACTTCAAT 1380
 TGAGAAATAC GATATATCAA AGAGTGATCC ATTGGCTGCA GATTCAAACG ATATCGAAAG 1440
 TGTCCCTTTG TCCAACCTGG AACTTGACTT TGATCTTAAT ATGACTGATG AACCCAGTGT 1500
 AAGCCCTTCG TTATTTGATG GATCTACTTC TGTTTCTCGT TCGTCCGCCA CGTCTATAAC 1560
 GTCACTAGGT GTAAAGAGGG CTTAATGAGT ATTTTATCTG CAATTACGGA TACGGTTGGT 1620
 CTTATGTAGA TACATATAAA TATATATCTT TTTCTTTCTT TTTCTTAGTC AGGATTGTCG 1680
 TTTAGCATAA TATACATGTA GTTTATTTAA TCACATACCA CTGATTATCT TTAGAATTTT 1740
 ATAAATTTTT GAAATAAATG GGTGGCTTTT AATGGTGTCT ATGTTAAGTG AGGCTTTTAG 1800
 AATGCTCTTC CTGCTTTGTT TATTATATGT GTATGAAAGA TATGTATGTA TTTACATGTG 1860
 TTTGTAGCGT CCCCAGTCAA AACCTGTGCG CTATACCTAA ATGGATTGAT AATCTTCATT 1920
 CACTAATTCT AAAATAGACT TCTTCCCCAA AGAACGGTGT AACGATGAGG CTCTATCCAG 1980
 CTGCTTATCT AAATCAACTT TAACGATGGA TGATCTTATG ACACGGGGAT CTTTCTTTAA 2040
 AGTTCTTAGA ATTTCAGACT GTACCGCAGC TGATGAATCA AACAGCATTA AAAAGTGATA 2100
 TGCTCGAAAA TGTTTTTCCT GGTCTTTCTT CATTATTTTA GGAAGATACC TTATGCCCCAT 2160
 GGGTACAATG TCCCTCACCA CACCTCTGTT TTGAATAATC AGTTTCCCGA TTGTGGAAGA 2220
 CAATTCTTTT GCTTCCAAC TTTGGCGCATT GGAGTTGGTT ATGCGAACAA GTCCGATCAG 2280
 CTCATAAAGC ATCTTAGTGA AAAGGGTGGT TTTGCGTTAT TCTTTCCTCT GTTGAAGCTT 2340

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5					10					15
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20					25					30
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35					40					45
Val	His	Tyr	Ile	Phe	Leu	Gly	Ser	Ile	Met	Leu	Phe	Val	Phe	Ile
				50					55					60
Thr	Asn	Pro	Ala	Pro	Trp	Ile	Phe	Lys	Ile	Leu	Phe	Tyr	Cys	Phe
				65					70					75
Leu	Gly	Thr	Leu	Phe	Ile	Ile	Pro	Ala	Thr	Ser	Gln	Phe	Phe	Phe
				80					85					90
Asn	Ala	Leu	Pro	Ile	Leu	Thr	Trp	Val	Ala	Leu	Tyr	Phe	Thr	Ser
				95					100					105
Ser	Tyr	Phe	Pro	Asp	Asp	Arg	Arg	Pro	Pro	Ile	Thr	Val	Lys	Val

110	115	120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp		
125	130	135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp		
140	145	150
Leu Pro Tyr Gly Leu Phe His Tyr Gly Ala Pro Phe Val Val Ala		
155	160	165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr		
170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln		
185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly		
200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu		
215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Ala		
230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His		
245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		
275	280	285

Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met		
	290	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
	305	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
	320	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
	335	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
	350	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
	365	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
	380	390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala		
	395	400

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2340 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
TTTCTTTCTG TCAAAGAATA ATAAAGTGCC CATCAGTGTT CATATTTGTT ACAAAGTGGT 60
TTTCTGATTT GGTACTACTG CAGAGGCGTA TTTTTTGCTT CAGTTACCAT AGCGTAAGAA 120
CACTAGCGAC TTTTGTTCGT GAACCAACAG AGTAGGATTT CTACTGCTAC ATCTCTTAGG 180
TAGTTGGTTA GTCCGATCGC TCACTTTTGG TTGTTGTTAA GTACTTCATA AGTTTATCCT 240
TTTCCTTTTT CACACTGAGC TACTTTGGGT ATAGCTTTTG GCCCAAGGAT CTTTGAATTT 300
TCTCCAAAAG TACTTTATTT TATATCCTAC AGGTTGCGGT TTCATATTT TAAAAAGCTT 360
TTTAATCATT CCTTTGCGTA TGGCAAACCC TTTTTCGAGA TGGTTTCTAT CAGAGAGACC 420
TCCAAACTGC CATGTAGCCG ATTTAGAAAC AAGTTTAGAT CCCCATCAAA CGTTGTTGAA 480
GGTGCAAAAA TACAAACCCG CTTTAAGCGA CTGGGTGCAT TACATCTTCT TGGGATCCAT 540
CATGCTGTTT GTGTTCATTA CTAATCCCGC ACCTTGGATC TTCAAGATCC TTTTTTATTG 600
TTTCTTGGGC ACTTTATTCA TCATTCCAGC TACGTCACAG TTTTCTTCA ATGCCTTGCC 660
CATCCTAACA TGGGTGGCGC TGTATTTTAC TTCATCGTAC TTTCCAGATG ACCGCAGGCC 720
TCCTATTACT GTCAAAGTGT TACCAGCGGT GGAAACAATT TTATACGGCG ACAATTTAAG 780
TGATATTCTT GCAACATCGA CGAATTCCTT TTTGGACATT TTAGCATGGT TACCGTACGG 840
ACTATTTTCAT TTTGGGGCCC CATTTGTCGT TGCTGCCATC TTATTCGTAT TTGGTCCACC 900
AACTGTTTTG CAAGGTTATG CTTTTGCATT TGGTTATATG AACCTGTTTG GTGTTATCAT 960
GCAAAATGTC TTTCCAGCCG CTCCCCCATG GTATAAAATT CTCTATGGAT TGCAATCAGC 1020
CAACTATGAT ATGCATGGCT CGCCTGGTGG ATTAGCTAGA ATTGATAAGC TACTCGGTAT 1080
TAATATGTAT ACTACAGCTT TTTCAAATTC CTCCGTCATT TTCGGTGCTT TTCCTTCACT 1140
GCATTCCGGG TGTGCTACTA TGGAAGCCCT GTTTTTCTGT TATTGTTTTT CAAAATTGAA 1200
```

GCCCTTGTTT ATTGCTTATG TTTGCTGGTT ATGGTGGTCA ACTATGTATC TGACACACCA 1260
 TTATTTTGTG GACCTTATGG CAGGTTCTGT GCTGTCATAC GTTATTTTCC AGTACACAAA 1320
 GTACACACAT TTACCAATTG TAGATACATC TCTTTTTTGC AGATGGTCAT ACACTTCAAT 1380
 TGAGAAATAC GATATATCAA AGAGTGATCC ATTGGCTGCA GATTCAAACG ATATCGAAAG 1440
 TGTCCCTTTG TCCAACCTGG AACTTGACTT TGATCTTAAT ATGACTGATG AACCCAGTGT 1500
 AAGCCCTTCG TTATTTGATG GATCTACTTC TGTTTCTCGT TCGTCCGCCA CGTCTATAAC 1560
 GTCACTAGGT GTAAAGAGGG CTTAATGAGT ATTTTATCTG CAATTACGGA TACGGTTGGT 1620
 CTTATGTAGA TACATATAAA TATATATCTT TTTCTTTCTT TTTCTTAGTC AGGATTGTCTG 1680
 TTTAGCATAA TATACATGTA GTTTATTTAA TCACATACCA CTGATTATCT TTAGAATTTT 1740
 ATAAATTTTT GAAATAAATG GGTGGCTTTT AATGGTGTCT ATGTTAAGTG AGGCTTTTAG 1800
 AATGCTCTTC CTGCTTTGTT TATTATATGT GTATGAAAGA TATGTATGTA TTTACATGTG 1860
 TTTGTAGCGT CCCCAGTCAA AACCTGTGCG CTATACCTAA ATGGATTGAT AATCTTCATT 1920
 CACTAATTCT AAAATAGACT TCTTCCCCAA AGAACGGTGT AACGATGAGG CTCTATCCAG 1980
 CTGCTTATCT AAATCAACTT TAACGATGGA TGATCTTATG ACACGGGGAT CTTTCTTTAA 2040
 AGTTCTTAGA ATTTCAGACT GTACCGCAGC TGATGAATCA AACAGCATTAA AAAAGTGATA 2100
 TGCTCGAAAA TGTTTTTCCT GGTCTTTCTT CATTATTTTA GGAAGATACC TTATGCCCAT 2160
 GGGTACAATG TCCCTCACCA CACCTCTGTT TTGAATAATC AGTTTCCCGA TTGTGGAAGA 2220
 CAATTCTTTT GCTTCCAAC TTTGGCGCATT GGAGTTGGTT ATGCGAACAA GTCCGATCAG 2280
 CTCATAAAGC ATCTTAGTGA AAAGGGTGGT TTTGCGTTAT TCTTTCCTCT GTTGAAGCTT 2340

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5					10					15
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20					25					30
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35					40					45
Val	His	Tyr	Ile	Phe	Leu	Gly	Ser	Ile	Met	Leu	Phe	Val	Phe	Ile
				50					55					60
Thr	Asn	Pro	Ala	Pro	Trp	Ile	Phe	Lys	Ile	Leu	Phe	Tyr	Cys	Phe
				65					70					75
Leu	Gly	Thr	Leu	Phe	Ile	Ile	Pro	Ala	Thr	Ser	Gln	Phe	Phe	Phe
				80					85					90
Asn	Ala	Leu	Pro	Ile	Leu	Thr	Trp	Val	Ala	Leu	Tyr	Phe	Thr	Ser
				95					100					105
Ser	Tyr	Phe	Pro	Asp	Asp	Arg	Arg	Pro	Pro	Ile	Thr	Val	Lys	Val
				110					115					120
Leu	Pro	Ala	Val	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asp

	125	130	135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp			
	140	145	150
Leu Pro Tyr Gly Leu Phe His Phe Gly Ala Pro Phe Val Val Ala			
	155	160	165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr			
	170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln			
	185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly			
	200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu			
	215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Ala			
	230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His			
	245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe			
	260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp			
	275	280	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met			
	290	295	300

Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
	305	310 315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
	320	325 330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
	335	340 345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
	350	355 360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
	365	370 375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
	380	385 390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala		
	395	400

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5340 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGCGCTTCTA	TTTTCCTCCC	CACCGCGAGG	CGGAAATGGC	ACATTTTTTTT	TCTTTTGCTT	60
CTGTGCTTTT	GCTGTAATTT	TTGGCATGTG	CTATTGTATG	AAGATAACGC	GTGGTTCCGT	120
GGAAATAGCC	GGAAATTTTG	CCGGGAATAT	GACGGACATG	ATTTAACACC	CGTGGAAATG	180
AAAAAAGCCA	AGGTAAGAAA	GTGGCAATAT	TTTTCCTACA	AATAGATCTG	CTGTCCCTTA	240
GATGATTACC	ATACATATAT	ATATTTATTA	CACACATCTG	TCAGAGGTAG	CTAGCGAAGG	300
TGTCACTGAA	ATATTTTTTG	TTCCAGTTAG	TATAAATACG	GAGGTAGAAC	AGCTCTCCGC	360
GTGTATATCT	TTTTTTGCGC	TATACAAGAA	CAGGAAGAAC	GCATTTCCAT	ACCTTTTTCT	420
CCTTACAGGT	GCCCTCTGAG	TAGTGTACG	AACGAGGAAA	AAGATTAATA	TTACTGTTTT	480
TATATTCAAA	AAGAGTAAAG	CCGTTGCTAT	ATACGAATAT	GACGATTACC	GTGGGGGATG	540
CAGTTTCGGA	GACGGAGCTG	GAAAACAAAA	GTCAAAACGT	GGTACTATCT	CCCAAGGCAT	600
CTGCTTCTTC	AGACATAAGC	ACAGATGTTG	ATAAAGACAC	ATCGTCTTCT	TGGGATGACA	660
AATCTTTGCT	GCCTACAGGT	GAATATATTG	TGGACAGAAA	TAAGCCCCAA	ACCTACTTGA	720
ATAGCGATGA	TATCGAAAAA	GTGACAGAAT	CTGATATTTT	CCCTCAGAAA	CGTCTGTTTT	780
CATTCTTGCA	CTCTAAGAAA	ATTCCAGAAG	TACCACAAAC	CGATGACGAG	AGGAAGATAT	840
ATCCTCTGTT	CCATACAAAT	ATTATCTCTA	ACATGTTTTT	TTGGTGGGTT	CTACCCATCC	900
TGCGAGTTGG	TTATAAGAGA	ACGATACAGC	CGAACGATCT	CTTCAAAATG	GATCCGAGGA	960
TGTCTATAGA	GACCCTTTAT	GACGACTTTG	AAAAAAACAT	GATTTACTAT	TTTGAGAAGA	1020
CGAGGAAAAA	ATACCGTAAA	AGACATCCAG	AAGCGACAGA	AGAAGAGGTT	ATGGAAAATG	1080
CCAAACTACC	TAAACATACA	GTTCTGAGAG	CTTTATTATT	CACTTTTAAG	AAACAGTACT	1140
TCATGTCGAT	AGTGTTTGCA	ATTCTCGCTA	ATTGTACATC	CGGTTTTAAC	CCCATGATTA	1200
CCAAGAGGCT	AATTGAGTTT	GTCGAAGAAA	AGGCTATTTT	TCATAGCATG	CATGTTAACA	1260
AAGGTATTGG	TTACGCTATT	GGTGCATGTT	TGATGATGTT	CGTTAACGGG	TTGACGTTCA	1320
ATCATTTCTT	TCATACATCC	CAACTGACTG	GTGTGCAAGC	TAAGTCTATT	CTTACTAAAG	1380

CTGCCATGAA GAAAATGTTT AATGCATCTA ATTATGCGAG ACATTGTTTT CCTAACGGTA 1440
 AAGTGACTTC TTTTGTAACA ACAGATCTCG CTAGAATTGA ATTTGCCTTA TCTTTTCAGC 1500
 CGTTTTTGGC TGGGTTCCTT GCAATTTTGG CTATTTGCAT TGTTTTATTG ATCGTTAACC 1560
 TTGGACCCAT TGCCTTAGTT GGGATTGGTA TTTTTTTCGG TGGGTTTTTC ATATCCTTAT 1620
 TTGCATTTAA GTTAATTCTG GGCTTTAGAA TTGCTGCGAA CATCTTCACT GATGCTAGAG 1680
 TTACCATGAT GAGAGAAGTG CTGAATAATA TAAAAATGAT TAAATATTAT ACGTGGGAGG 1740
 ATGCGTATGA AAAAAATATT CAAGATATTA GGACCAAAGA GATTTCTAAA GTTAGAAAAA 1800
 TGCAACTATC AAGAAATTTT TTGATTGCTA TGGCCATGTC TTTGCCTAGT ATTGCTTCAT 1860
 TGGTCACTTT CCTTGCAATG TACAAAGTTA ATAAAGGAGG CAGGCAACCT GGTAATATTT 1920
 TTGCCTCTTT ATCTTTATTT CAGGTCTTGA GTTTGCAAAT GTTTTTCTTA CCTATTGCTA 1980
 TTGGTACTGG AATTGACATG ATCATTGGAT TGGGCCGTTT GCAAAGCTTA TTGGAGGCTC 2040
 CAGAAGATGA TCCAAATCAG ATGATTGAAA TGAAGCCCTC TCCTGGCTTT GATCCAAAAT 2100
 TGGCTCTAAA AATGACACAT TGCTCATTTG AGTGGGAAGA TTATGAATTA AACGACGCTA 2160
 TTGAAGAAGC AAAAGGAGAA GCTAAAGATG AAGGTAAAAA GAACAAAAAA AAGCGTAAGG 2220
 ATACATGGGG TAAGCCATCT GCAAGTACTA ATAAGGCGAA AAGATTGGAC AATATGTTGA 2280
 AAGACAGAGA CGGCCCGGAA GATTTAGAAA AAACCTTCGT TAGGGGTTTC AAGGACTTGA 2340
 ACTTCGATAT TAAAAAGGGC GAATTTATTA TGATTACGGG ACCTATTGGT ACTGGTAAAT 2400
 CTTCAATTATT GAATGCGATG GCAGGATCAA TGAGAAAAAT TGATGGTAAG GTTGAAGTCA 2460
 ACGGGGACTT ATTAATGTGT GGTATCCAT GGATTCAAAA TGCATCTGTA AGAGATAACA 2520
 TCATATTCGG TTCACCATT C AATAAAGAAA AGTATGATGA AGTAGTTCGT GTTTGCTCTT 2580
 TGAAAGCTGA TCTGGATATT TTACCGGCAG GCGATATGAC CGAAATTGGG GAACGTGGTA 2640
 TTA CTTTATC TGGTGGTCAA AAGGCACGTA TCAATTTAGC CAGGTCTGTT TATAAGAAGA 2700
 AGGATATTTA TGTATTCGAC GATGTCCTAA GTGCTGTCGA TTCTCGTGTT GGTAACACA 2760

TCATGGATGA ATGTCTAACC GGAATGCTTG CTAATAAAAC CAGAATTTTA GCAACGCATC 2820
 AGTTGTCACT GATTGAGAGA GCTTCTAGAG TCATCGTTTT AGGTACTGAT GGCCAAGTCG 2880
 ATATTGGTAC TGTTGATGAG CTAAAAGCTC GTAATCAAAC TTTGATAAAT CTTTTACAAT 2940
 TCTCTTCTCA AAATTCGGAG AAAGAGGATG AAGAACAGGA AGCGGTTGTT TCCGGTGAAT 3000
 TGGGACAACT AAAATATGAA CCAGAGGTAA AGGAATTGAC TGAAGTGAAG AAAAAGGCTA 3060
 CAGAAATGTC ACAAAGTACA AATAGTGGTA AAATTGTAGC GGATGGTCAT ACTAGTAGTA 3120
 AAGAAGAAAG AGCAGTCAAT AGTATCAGTC TGAAAATATA CCGTGAATAC ATTAAAGCTG 3180
 CAGTAGGTAA GTGGGGTTTT ATCGCACTAC CGTTGTATGC AATTTTAGTC GTTGAACCA 3240
 CATTCTGCTC ACTTTTTTCT TCCGTTTGGT TATCTTACTG GACTGAGAAT AAATTCAAAA 3300
 ACAGACCACC CAGTTTTTAT ATGGGTCTTT ACTCCTTCTT TGTGTTTGCT GCTTTCATAT 3360
 TCATGAATGG CCAGTTCACC ATACTTTGCG CAATGGGTAT TATGGCATCG AAATGGTTAA 3420
 ATTTGAGGGC TGTGAAAAGA ATTTTACACA CTCCAATGTC ATACATAGAT ACCACACCTT 3480
 TGGGACGTAT TCTGAACAGA TTCACAAAAG ATACAGATAG CTTAGATAAT GAGTTAACCG 3540
 AAAGTTTACG GTTGATGACA TCTCAATTTG CTAATATTGT AGGTGTTTGC GTCATGTGTA 3600
 TTGTTTACTT GCCGTGGTTT GCTATCGCAA TTCCGTTTCT TTTGGTCATC TTTGTTCTGA 3660
 TTGCTGATCA TTATCAGAGT TCTGGTAGAG AAATTAAAAG ACTTGAAGCT GTGCAACGGT 3720
 CTTTTGTTTA CAATAATTTA AATGAAGTTT TGGGTGGGAT GGATACAATC AAAGCATACC 3780
 GAAGTCAGGA ACGATTTTTG GCGAAATCAG ATTTTTTGAT CAACAAGATG AATGAGGCGG 3840
 GATACCTTGT AGTTGTCCTG CAAAGATGGG TAGGTATTTT CCTTGATATG GTTGCTATCG 3900
 CATTTGCACT AATTATTACG TTATTGTGTG TTACGAGAGC CTTTCCTATT TCCGCGGCTT 3960
 CAGTTGGTGT TTTGTTGACT TATGTATTAC AATTGCCTGG TCTATTAAAT ACCATTTTAA 4020
 GGGCAATGAC TCAAACAGAG AATGACATGA ATAGTGCCGA AAGATTGGTA ACATATGCAA 4080
 CTGAACTACC ACTAGAGGCA TCCTATAGAA AGCCCGAAAT GACACCTCCA GAGTCATGGC 4140

CCTCAATGGG CGAAATAATT TTTGAAAATG TTGATTTTGC CTATAGACCT GGTTTACCTA 4200
TAGTTTTAAA AAATCTTAAC TTGAATATCA AGAGTGGGGA AAAAATTGGT ATCTGTGGTC 4260
GTACAGGTGC TGGTAAGTCC ACTATTATGA GTGCCCTTTA CAGGTTGAAT GAATTGACCG 4320
CAGGTAAAAT TTTAATTGAC AATGTTGATA TAAGTCAGCT GGGACTTTTC GATTTAAGAA 4380
GAAAATTAGC CATCATTTCA CAAGATCCAG TATTATTTAG GGGTACGATT CGCAAGAACT 4440
TAGATCCATT TAATGAGCGT ACAGATGACG AATTATGGGA TGCATTGGTG AGAGGTGGTG 4500
CTATCGCCAA GGATGACTTG CCGGAAGTGA AATTGCAAAA ACCTGATGAA AATGGTACTC 4560
ATGGTAAAAT GCATAAGTTC CATTTAGATC AAGCAGTGGA AGAAGAGGGC TCCAATTTCT 4620
CCTTAGGTGA GAGACAACTA TTAGCATTAA CAAGGGCATT GTCCGCCAA TCAAAAATAT 4680
TGATTTTGGA TGAGGCTACA TCCTCAGTGG ACTACGAAAC GGATGGCAAA ATCCAAACAC 4740
GTATTGTTGA GGAATTTGGA GATTGTACAA TTTTGTGTAT TGCTCACAGA CTGAAGACCA 4800
TTGTAAATTA TGATCGTATT CTTGTTTTAG AGAAGGGTGA AGTCGCAGAA TTCGATACAC 4860
CATGGACGTT GTTTAGTCAA GAAGATAGTA TTTTCAGAAG CATGTGTTCT AGATCTGGTA 4920
TTGTGGAAAA TGATTTTCGAG AACAGAAGTT AATTTATATT ATTTGTTGCA TGATTTTTCT 4980
CTTTTATTTA TTTATATGTT GCCGATGGTA CAAATTAGTA CTAGAAAAGA AAACCCACTA 5040
CTATGACTTG CAGAAAAAGT TATGTGTGCC ATAGATAGAT ATAATTGCAT ACCCACATCG 5100
TATACTCAA ATTCCGAAAA GAACATTTCA TTTTTTATGA GGCAAACCTGA ACAACGCTTC 5160
GGTCCTTTTT TCATTCTAGA AATATATATT TATACATCAT TTTCAGAAGA TATTCAAAGA 5220
ACTTATTGGG ATGTCTATTT ACTGAATAAA GTATACACAA AAAACGAATT TAAAATGGAA 5280
GGCATAAATA GAAAACTTAG AAGTGAAAAT CCTAAAACCG AAGGATATTT CAAATACGTA 5340

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1477 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met	Thr	Ile	Thr	Val	Gly	Asp	Ala	Val	Ser	Glu	Thr	Glu	Leu	Glu
				5					10					15
Asn	Lys	Ser	Gln	Asn	Val	Val	Leu	Ser	Pro	Lys	Ala	Ser	Ala	Ser
				20					25					30
Ser	Asp	Ile	Ser	Thr	Asp	Val	Asp	Lys	Asp	Thr	Ser	Ser	Ser	Trp
				35					40					45
Asp	Asp	Lys	Ser	Leu	Leu	Pro	Thr	Gly	Glu	Tyr	Ile	Val	Asp	Arg
				50					55					60
Asn	Lys	Pro	Gln	Thr	Tyr	Leu	Asn	Ser	Asp	Asp	Ile	Glu	Lys	Val
				65					70					75
Thr	Glu	Ser	Asp	Ile	Phe	Pro	Gln	Lys	Arg	Leu	Phe	Ser	Phe	Leu
				80					85					90
His	Ser	Lys	Lys	Ile	Pro	Glu	Val	Pro	Gln	Thr	Asp	Asp	Glu	Arg
				95					100					105
Lys	Ile	Tyr	Pro	Leu	Phe	His	Thr	Asn	Ile	Ile	Ser	Asn	Met	Phe
				110					115					120

Phe	Trp	Trp	Val	Leu	Pro	Ile	Leu	Arg	Val	Gly	Tyr	Lys	Arg	Thr
				125					130					135
Ile	Gln	Pro	Asn	Asp	Leu	Phe	Lys	Met	Asp	Pro	Arg	Met	Ser	Ile
				140					145					150
Glu	Thr	Leu	Tyr	Asp	Asp	Phe	Glu	Lys	Asn	Met	Ile	Tyr	Tyr	Phe
				155					160					165
Glu	Lys	Thr	Arg	Lys	Lys	Tyr	Arg	Lys	Arg	His	Pro	Glu	Ala	Thr
				170					175					180
Glu	Glu	Glu	Val	Met	Glu	Asn	Ala	Lys	Leu	Pro	Lys	His	Thr	Val
				185					190					195
Leu	Arg	Ala	Leu	Leu	Phe	Thr	Phe	Lys	Lys	Gln	Tyr	Phe	Met	Ser
				200					205					210
Ile	Val	Phe	Ala	Ile	Leu	Ala	Asn	Cys	Thr	Ser	Gly	Phe	Asn	Pro
				215					220					225
Met	Ile	Thr	Lys	Arg	Leu	Ile	Glu	Phe	Val	Glu	Glu	Lys	Ala	Ile
				230					235					240
Phe	His	Ser	Met	His	Val	Asn	Lys	Gly	Ile	Gly	Tyr	Ala	Ile	Gly
				245					250					255
Ala	Cys	Leu	Met	Met	Phe	Val	Asn	Gly	Leu	Thr	Phe	Asn	His	Phe
				260					265					270
Phe	His	Thr	Ser	Gln	Leu	Thr	Gly	Val	Gln	Ala	Lys	Ser	Ile	Leu
				275					280					285
Thr	Lys	Ala	Ala	Met	Lys	Lys	Met	Phe	Asn	Ala	Ser	Asn	Tyr	Ala

	290		295		300
Arg His Cys Phe	Pro Asn Gly Lys Val	Thr Ser Phe Val	Thr Thr		
	305		310		315
Asp Leu Ala Arg	Ile Glu Phe Ala Leu	Ser Phe Gln Pro	Phe Leu		
	320		325		330
Ala Gly Phe Pro	Ala Ile Leu Ala Ile	Cys Ile Val Leu	Leu Ile		
	335		340		345
Val Asn Leu Gly	Pro Ile Ala Leu Val	Gly Ile Gly Ile	Phe Phe		
	350		355		360
Gly Gly Phe Phe	Ile Ser Leu Phe Ala	Phe Lys Leu Ile	Leu Gly		
	365		370		375
Phe Arg Ile Ala	Ala Asn Ile Phe Thr	Asp Ala Arg Val	Thr Met		
	380		385		390
Met Arg Glu Val	Leu Asn Asn Ile Lys	Met Ile Lys Tyr	Tyr Thr		
	395		400		405
Trp Glu Asp Ala	Tyr Glu Lys Asn Ile	Gln Asp Ile Arg	Thr Lys		
	410		415		420
Glu Ile Ser Lys	Val Arg Lys Met Gln	Leu Ser Arg Asn	Phe Leu		
	425		430		435
Ile Ala Met Ala	Met Ser Leu Pro Ser	Ile Ala Ser Leu	Val Thr		
	440		445		450
Phe Leu Ala Met	Tyr Lys Val Asn Lys	Gly Gly Arg Gln	Pro Gly		
	455		460		465

Asn	Ile	Phe	Ala	Ser	Leu	Ser	Leu	Phe	Gln	Val	Leu	Ser	Leu	Gln
				470					475					480
Met	Phe	Phe	Leu	Pro	Ile	Ala	Ile	Gly	Thr	Gly	Ile	Asp	Met	Ile
				485					490					495
Ile	Gly	Leu	Gly	Arg	Leu	Gln	Ser	Leu	Leu	Glu	Ala	Pro	Glu	Asp
				500					505					510
Asp	Pro	Asn	Gln	Met	Ile	Glu	Met	Lys	Pro	Ser	Pro	Gly	Phe	Asp
				515					520					525
Pro	Lys	Leu	Ala	Leu	Lys	Met	Thr	His	Cys	Ser	Phe	Glu	Trp	Glu
				530					535					540
Asp	Tyr	Glu	Leu	Asn	Asp	Ala	Ile	Glu	Glu	Ala	Lys	Gly	Glu	Ala
				545					550					555
Lys	Asp	Glu	Gly	Lys	Lys	Asn	Lys	Lys	Lys	Arg	Lys	Asp	Thr	Trp
				560					565					570
Gly	Lys	Pro	Ser	Ala	Ser	Thr	Asn	Lys	Ala	Lys	Arg	Leu	Asp	Asn
				575					580					585
Met	Leu	Lys	Asp	Arg	Asp	Gly	Pro	Glu	Asp	Leu	Glu	Lys	Thr	Ser
				590					595					600
Phe	Arg	Gly	Phe	Lys	Asp	Leu	Asn	Phe	Asp	Ile	Lys	Lys	Gly	Glu
				605					610					615
Phe	Ile	Met	Ile	Thr	Gly	Pro	Ile	Gly	Thr	Gly	Lys	Ser	Ser	Leu
				620					625					630
Leu	Asn	Ala	Met	Ala	Gly	Ser	Met	Arg	Lys	Ile	Asp	Gly	Lys	Val

635	640	645
Glu Val Asn Gly Asp Leu Leu Met Cys Gly Tyr Pro Trp Ile Gln		
650	655	660
Asn Ala Ser Val Arg Asp Asn Ile Ile Phe Gly Ser Pro Phe Asn		
665	670	675
Lys Glu Lys Tyr Asp Glu Val Val Arg Val Cys Ser Leu Lys Ala		
680	685	690
Asp Leu Asp Ile Leu Pro Ala Gly Asp Met Thr Glu Ile Gly Glu		
695	700	705
Arg Gly Ile Thr Leu Ser Gly Gly Gln Lys Ala Arg Ile Asn Leu		
710	715	720
Ala Arg Ser Val Tyr Lys Lys Lys Asp Ile Tyr Val Phe Asp Asp		
725	730	735
Val Leu Ser Ala Val Asp Ser Arg Val Gly Lys His Ile Met Asp		
740	745	750
Glu Cys Leu Thr Gly Met Leu Ala Asn Lys Thr Arg Ile Leu Ala		
755	760	765
Thr His Gln Leu Ser Leu Ile Glu Arg Ala Ser Arg Val Ile Val		
770	775	780
Leu Gly Thr Asp Gly Gln Val Asp Ile Gly Thr Val Asp Glu Leu		
785	790	795
Lys Ala Arg Asn Gln Thr Leu Ile Asn Leu Leu Gln Phe Ser Ser		
800	805	810

Gln	Asn	Ser	Glu	Lys	Glu	Asp	Glu	Glu	Gln	Glu	Ala	Val	Val	Ser
				815					820					825
Gly	Glu	Leu	Gly	Gln	Leu	Lys	Tyr	Glu	Pro	Glu	Val	Lys	Glu	Leu
				830					835					840
Thr	Glu	Leu	Lys	Lys	Lys	Ala	Thr	Glu	Met	Ser	Gln	Thr	Ala	Asn
				845					850					855
Ser	Gly	Lys	Ile	Val	Ala	Asp	Gly	His	Thr	Ser	Ser	Lys	Glu	Glu
				860					865					870
Arg	Ala	Val	Asn	Ser	Ile	Ser	Leu	Lys	Ile	Tyr	Arg	Glu	Tyr	Ile
				875					880					885
Lys	Ala	Ala	Val	Gly	Lys	Trp	Gly	Phe	Ile	Ala	Leu	Pro	Leu	Tyr
				890					895					900
Ala	Ile	Leu	Val	Val	Gly	Thr	Thr	Phe	Cys	Ser	Leu	Phe	Ser	Ser
				905					910					915
Val	Trp	Leu	Ser	Tyr	Trp	Thr	Glu	Asn	Lys	Phe	Lys	Asn	Arg	Pro
				920					925					930
Pro	Ser	Phe	Tyr	Met	Gly	Leu	Tyr	Ser	Phe	Phe	Val	Phe	Ala	Ala
				935					940					945
Phe	Ile	Phe	Met	Asn	Gly	Gln	Phe	Thr	Ile	Leu	Cys	Ala	Met	Gly
				950					955					960
Ile	Met	Ala	Ser	Lys	Trp	Leu	Asn	Leu	Arg	Ala	Val	Lys	Arg	Ile
				965					970					975
Leu	His	Thr	Pro	Met	Ser	Tyr	Ile	Asp	Thr	Thr	Pro	Leu	Gly	Arg

980	985	990
Ile Leu Asn Arg Phe Thr Lys Asp Thr Asp Ser Leu Asp Asn Glu		
995	1000	1005
Leu Thr Glu Ser Leu Arg Leu Met Thr Ser Gln Phe Ala Asn Ile		
1010	1015	1020
Val Gly Val Cys Val Met Cys Ile Val Tyr Leu Pro Trp Phe Ala		
1025	1030	1035
Ile Ala Ile Pro Phe Leu Leu Val Ile Phe Val Leu Ile Ala Asp		
1040	1045	1050
His Tyr Gln Ser Ser Gly Arg Glu Ile Lys Arg Leu Glu Ala Val		
1055	1060	1065
Gln Arg Ser Phe Val Tyr Asn Asn Leu Asn Glu Val Leu Gly Gly		
1070	1075	1080
Met Asp Thr Ile Lys Ala Tyr Arg Ser Gln Glu Arg Phe Leu Ala		
1085	1090	1095
Lys Ser Asp Phe Leu Ile Asn Lys Met Asn Glu Ala Gly Tyr Leu		
1100	1105	1110
Val Val Val Leu Gln Arg Trp Val Gly Ile Phe Leu Asp Met Val		
1115	1120	1125
Ala Ile Ala Phe Ala Leu Ile Ile Thr Leu Leu Cys Val Thr Arg		
1130	1135	1140
Ala Phe Pro Ile Ser Ala Ala Ser Val Gly Val Leu Leu Thr Tyr		
1145	1150	1155

Val Leu Gln Leu Pro Gly Leu Leu Asn Thr Ile Leu Arg Ala Met			
	1160	1165	1170
Thr Gln Thr Glu Asn Asp Met Asn Ser Ala Glu Arg Leu Val Thr			
	1175	1180	1185
Tyr Ala Thr Glu Leu Pro Leu Glu Ala Ser Tyr Arg Lys Pro Glu			
	1190	1195	1200
Met Thr Pro Pro Glu Ser Trp Pro Ser Met Gly Glu Ile Ile Phe			
	1205	1210	1215
Glu Asn Val Asp Phe Ala Tyr Arg Pro Gly Leu Pro Ile Val Leu			
	1220	1225	1230
Lys Asn Leu Asn Leu Asn Ile Lys Ser Gly Glu Lys Ile Gly Ile			
	1235	1240	1245
Cys Gly Arg Thr Gly Ala Gly Lys Ser Thr Ile Met Ser Ala Leu			
	1250	1255	1260
Tyr Arg Leu Asn Glu Leu Thr Ala Gly Lys Ile Leu Ile Asp Asn			
	1265	1270	1275
Val Asp Ile Ser Gln Leu Gly Leu Phe Asp Leu Arg Arg Lys Leu			
	1280	1285	1290
Ala Ile Ile Pro Gln Asp Pro Val Leu Phe Arg Gly Thr Ile Arg			
	1295	1300	1305
Lys Asn Leu Asp Pro Phe Asn Glu Arg Thr Asp Asp Glu Leu Trp			
	1310	1315	1320
Asp Ala Leu Val Arg Gly Gly Ala Ile Ala Lys Asp Asp Leu Pro			

1325	1330	1335
Glu Val Lys Leu Gln Lys Pro Asp Glu Asn Gly Thr His Gly Lys		
1340	1345	1350
Met His Lys Phe His Leu Asp Gln Ala Val Glu Glu Glu Gly Ser		
1355	1360	1365
Asn Phe Ser Leu Gly Glu Arg Gln Leu Leu Ala Leu Thr Arg Ala		
1370	1375	1380
Leu Val Arg Gln Ser Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser		
1385	1390	1395
Ser Val Asp Tyr Glu Thr Asp Gly Lys Ile Gln Thr Arg Ile Val		
1400	1405	1410
Glu Glu Phe Gly Asp Cys Thr Ile Leu Cys Ile Ala His Arg Leu		
1415	1420	1425
Lys Thr Ile Val Asn Tyr Asp Arg Ile Leu Val Leu Glu Lys Gly		
1430	1435	1440
Glu Val Ala Glu Phe Asp Thr Pro Trp Thr Leu Phe Ser Gln Glu		
1445	1450	1455
Asp Ser Ile Phe Arg Ser Met Cys Ser Arg Ser Gly Ile Val Glu		
1460	1465	1470
Asn Asp Phe Glu Asn Arg Ser		
1475		

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 bases
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTTGGTTAYA TGAAYTNTT YGGNGT 26

- (2) INFORMATION FOR SEQ ID NO:26:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 bases
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCTACAAART ARTGGTGNGT NARRTACAT 29

- (2) INFORMATION FOR SEQ ID NO:27:
 - (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2274 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: genomic DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

TTATATATAT TATTGATTTG TTCCTGTTGT TATTTAGTTT AGAATCAGAC GACTACACCA    60
GAACCACAAT TCAACCAACA CTTATATAGA ACCTGGCTTG GAAAAAAGTA ACATTTATCA    120
TTCCTATACT TTTTtagCAA ACATAATCCG TGTTTTACAT ATATTATTCA CCCAATATCA    180
TAACAAAAAC AAActGAATA ATGGCGTCTT CTATTTTGCG TTCCAAAATA ATACAAAAAC    240
CGTACCAATT ATTCCACTAC TATTTTCTTC TGGAGAAGGC TCCTGGTTCT ACAGTTAGTG    300
ATTTGAATTT TGATACAAAC ATACAAACGA GTTTACGTAA ATTAAAGCAT CATCATTGGA    360
CGGTGGGAGA AATATTCCAT TATGGGTTTT TGGTTTCCAT ACTTTTTTTC GTGTTTGTGG    420
TTTTCCCAGC TTCATTTTTT ATAAAATTAC CAATAATCTT AGCATTTGCT ACTTGTTTTT    480
TAATACCCTT AACATCACAA TTTTTTCTTC CTGCCTTGCC CGTTTTCACT TGGTTGGCAT    540
TATATTTTAC GTGTGCTAAA ATACCTCAAG AATGGAAACC AGCTATCACA GTTAAAGTTT    600
TACCAGCTAT GGAAACAATT TTGTACGGCG ATAATTTATC AAATGTTTTG GCAACCATCA    660
CTACCGGAGT GTTAGATATA TTGGCATGGT TACCATATGG GATTATTCAT TTCAGTTTCC    720
CATTTGTACT TGCTGCTATT ATATTTTAT TTGGGCCACC GACGGCATTa AGATCATTTG    780
GATTGTCCTT TGGTTATATG AACTTGCTTG GAGTCTTGAT TCAAATGGCA TTCCCAGCTG    840
CTCCTCCATG GTACAAAAAC TTGCACGGAT TAGAACCAGC TAATTATTCA ATGCACGGGT    900
CTCCTGGTGG ACTTGGAAGG ATAGATAAAT TGTTAGGTGT TGATATGTAT ACCACAGGGT    960

```

TTTCCAATTC ATCAATCATT TTTGGGGCAT TCCCATCGTT ACATTCAGGA TGTTGTATCA 1020
 TGGAAGTGTT ATTTTTGTGT TGGTTGTTTC CACGATTCAA GTTTGTGTGG GTTACATACG 1080
 CATCTTGGCT TTGGTGGAGC ACGATGTATT TGACCCATCA CTA CTTTGTGTC GATTTGATTG 1140
 GTGGAGCCAT GCTATCTTTG ACTGTTTTTG AGTTCACCAA ATATAAATAT TTGCCAAAAA 1200
 ACAAAGAAGG CCTTTTCTGT CGTTGGTCAT AACTGAAAT TGAAAAAATC GATATCCAAG 1260
 AGATTGACCC TTTATCATAC AATTATATCC CTGTCAACAG CAATGATAAT GAAAGCAGAT 1320
 TGTATACGAG AGTGTAACAA GAGTCTCAGG TTAGTCCCCC ACAGAGAGCT GAAACACCTG 1380
 AAGCATTTGA GATGTCAAAT TTTTCTAGGT CTAGACAAAG CTCAAAGACT CAGGTTCCAT 1440
 TGAGTAATCT TACTAACAAT GATCAAGTGT CTGGAATTAA CGAAGAGGAT GAAGAAGAAG 1500
 AAGGCGATGA AATTTCATCG AGTACTCCTT CGGTGTTTGA AGACGAACCA CAGGGTAGCA 1560
 CATATGCTGC ATCCTCAGCT ACATCAGTAG ATGATTTGGA TTCCAAAAGA AATTAGTAAA 1620
 ATAACAGTTT CTATTAATTT CTTTATTTCC TCCTAATTAA TGATTTTATG CTCAATACCT 1680
 AACTATCTG TTTTAAATTT CCTACTTTTT TTTTATTATT GTTGAGTTCA TTTGCTGTTT 1740
 ATTGAATATT TACAATTTTG CATTAATTAC CATCAATATA GAATGGGCAC AGTTTTTTTT 1800
 AGTTTTTTTTG TTTTGTGTT TGTCTTCTT TTTTACATT AATGTGTTTG GATTGTTTTA 1860
 GGTTCCTTTA TCCCTTAGCC CCCTCAGAAT ACTATTTTAT CTAATTAATT TGTTTTTTATT 1920
 TTCTGATATT TACCAATTGC TTTTCTTTT GGATATTTAT AATAGCATCC CCTAATAATT 1980
 AATATACAAC TGTTTCATAT ATATACGTGT ATGTCCTGTA GTGGTGGAAG CTGGAGTCAA 2040
 CATTTGTATT AATGTGTACA AGAAAGCAGT GTTAATGCTA CTATTATAAT TTTTGAGGTG 2100
 CAAATCAAGA GGTGAGCAGC TTTCTTATGG CTATGACCGT GAATGAAGGC TTGTAAACCA 2160
 CGTAATAAAC AAAAGCCAAC AAGTTTTTTT AGAGCCTTTA ACAACATACG CAATGAGAGT 2220
 GATTGCAATA CTACAAGATA TAGCCCAAAA AATTGAATGC ATTTCAACAA CAAC 2274

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met	Ala	Ser	Ser	Ile	Leu	Arg	Ser	Lys	Ile	Ile	Gln	Lys	Pro	Tyr	
				5				10						15	
Gln	Leu	Phe	His	Tyr	Tyr	Phe	Leu	Ser	Glu	Lys	Ala	Pro	Gly	Ser	
				20				25						30	
Thr	Val	Ser	Asp	Leu	Asn	Phe	Asp	Thr	Asn	Ile	Gln	Thr	Ser	Leu	
				35				40						45	
Arg	Lys	Leu	Lys	His	His	His	Trp	Thr	Val	Gly	Glu	Ile	Phe	His	
				50				55						60	
Tyr	Gly	Phe	Leu	Val	Ser	Ile	Leu	Phe	Phe	Val	Phe	Val	Val	Phe	
				65				70						75	
Pro	Ala	Ser	Phe	Phe	Ile	Lys	Leu	Pro	Ile	Ile	Leu	Ala	Phe	Ala	
				80				85						90	
Thr	Cys	Phe	Leu	Ile	Pro	Leu	Thr	Ser	Gln	Phe	Phe	Leu	Pro	Ala	
				95				100						105	

Leu	Pro	Val	Phe	Thr	Trp	Leu	Ala	Leu	Tyr	Phe	Thr	Cys	Ala	Lys
				110					115					120
Ile	Pro	Gln	Glu	Trp	Lys	Pro	Ala	Ile	Thr	Val	Lys	Val	Leu	Pro
				125					130					135
Ala	Met	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asn	Val	Leu
				140					145					150
Ala	Thr	Ile	Thr	Thr	Gly	Val	Leu	Asp	Ile	Leu	Ala	Trp	Leu	Pro
				155					160					165
Tyr	Gly	Ile	Ile	His	Phe	Ser	Phe	Pro	Phe	Val	Leu	Ala	Ala	Ile
				170					175					180
Ile	Phe	Leu	Phe	Gly	Pro	Pro	Thr	Ala	Leu	Arg	Ser	Phe	Gly	Phe
				185					190					195
Ala	Phe	Gly	Tyr	Met	Asn	Leu	Leu	Gly	Val	Leu	Ile	Gln	Met	Ala
				200					205					210
Phe	Pro	Ala	Ala	Pro	Pro	Trp	Tyr	Lys	Asn	Leu	His	Gly	Leu	Glu
				215					220					225
Pro	Ala	Asn	Tyr	Ser	Met	His	Gly	Ser	Pro	Gly	Gly	Leu	Gly	Arg
				230					235					240
Ile	Asp	Lys	Leu	Leu	Gly	Val	Asp	Met	Tyr	Thr	Thr	Gly	Phe	Ser
				245					250					255
Asn	Ser	Ser	Ile	Ile	Phe	Gly	Ala	Phe	Pro	Ser	Leu	His	Ser	Gly
				260					265					270
Cys	Cys	Ile	Met	Glu	Val	Leu	Phe	Leu	Cys	Trp	Leu	Phe	Pro	Arg

275	280	285
Phe Lys Phe Val Trp Val Thr Tyr Ala Ser Trp Leu Trp Trp Ser		
290	295	300
Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Ile Gly Gly		
305	310	315
Ala Met Leu Ser Leu Thr Val Phe Glu Phe Thr Lys Tyr Lys Tyr		
320	325	330
Leu Pro Lys Asn Lys Glu Gly Leu Phe Cys Arg Trp Ser Tyr Thr		
335	340	345
Glu Ile Glu Lys Ile Asp Ile Gln Glu Ile Asp Pro Leu Ser Tyr		
350	355	360
Asn Tyr Ile Pro Val Asn Ser Asn Asp Asn Glu Ser Arg Leu Tyr		
365	370	375
Thr Arg Val Tyr Gln Glu Ser Gln Val Ser Pro Pro Gln Arg Ala		
380	385	390
Glu Thr Pro Glu Ala Phe Glu Met Ser Asn Phe Ser Arg Ser Arg		
395	400	405
Gln Ser Ser Lys Thr Gln Val Pro Leu Ser Asn Leu Thr Asn Asn		
410	415	420
Asp Gln Val Ser Gly Ile Asn Glu Glu Asp Glu Glu Glu Glu Gly		
425	430	435
Asp Glu Ile Ser Ser Ser Thr Pro Ser Val Phe Glu Asp Glu Pro		
440	445	450

Gln Gly Ser Thr Tyr Ala Ala Ser Ser Ala Thr Ser Val Asp Asp

455

460

465

Leu Asp Ser Lys Arg Asn

470

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 243 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
TTTGAAAAAT TTGAATTTTA AAATTAATCC AATGGAAAAA ATTGGTATTT GTGGAAGAAC    60
CGGTGCTGGT AAATCATCAA TTATGACAGC ATTATATCGA TTATCAGAAT TAGAACTGGG    120
GAAAATTATT ATTGATGATA TTGATATTTT AACTTTGGGT TTAAAAGATC TTCGATCAAA    180
ATTATCAATT ATTCCTCAAG ATCCAGTATT ATTCCGAGGT TCAATTCGGA AAAACTTGGA    240
TCC                                                                    243
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Leu	Lys	Asn	Leu	Asn	Phe	Lys	Ile	Asn	Pro	Met	Glu	Lys	Ile	Gly
				5					10					15
Ile	Cys	Gly	Arg	Thr	Gly	Ala	Gly	Lys	Ser	Ser	Ile	Met	Thr	Ala
				20					25					30
Leu	Tyr	Arg	Leu	Ser	Glu	Leu	Glu	Leu	Gly	Lys	Ile	Ile	Ile	Asp
				35					40					45
Asp	Ile	Asp	Ile	Ser	Thr	Leu	Gly	Leu	Lys	Asp	Leu	Arg	Ser	Lys
				50					55					60
Leu	Ser	Ile	Ile	Pro	Gln	Asp	Pro	Val	Leu	Phe	Arg	Gly	Ser	Ile
				65					70					75
Arg	Lys	Asn	Leu	Asp										
				80										

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1601 bases
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: genomic DNA
 (iv) ANTI-SENSE: Yes
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AGGAAGATGA	CTTGCATCAA	AGATGGAGGA	AGTGGTACTG	GCAGGACGAT	CAATCAAATC	60
AGCAGCAGGA	CTAGGTAACG	GCTCAGGTGA	TGATGAACCC	ACGGACCATT	CATGATCGGT	120
GTTAGCAAGT	TCCATATTGT	TAAGACCACT	CATGAAGGCT	ACTGCATTAG	GGTTTTGAGT	180
AAAAGAATCC	CTTCCAAGTA	AGTATGGGCT	GCCGGTACGA	GCCAAGGAGT	TGCTGGTTTT	240
TTCGGAAAGA	CCATGACCGT	GGATAACAAA	CTCGTATTCC	CAACGAAGGA	TTTTACCAGT	300
TTGCAACTGT	GGGAGGCGTA	GCTTTTGAGC	AAAAACGAAG	CATATAATAG	CTAAACACAT	360
ACCGCCGACC	AAATCTACAA	AGTAGTGGTG	GGTAAGGTAC	ATAGTACACC	AGCAAAGCCA	420
TAGAACATAT	CCATAAAAGC	AGAAGCGGTA	TCGAGGAAAC	ACATGCGAAA	GGAAAAGTGC	480
TTCCAGCATG	GCCCATCCAG	CGTGAAGAGA	TGGAAAGGCA	CCAAAAACAA	CCGGAGAGTT	540
AGAAAAACCA	TCAGTGTAAG	TGCTAGTGCC	GAAGAGAGCA	TCAATACGGG	CCAATCCACC	600
AGGAGAGCCA	CGTACTGCAT	ACGTGGCAGG	TTCTAAACCA	TACATATTTT	CATACCAAGG	660
AGGAGAACAG	GGGAAAGCCA	TTTGGATAAG	AACACCAAAT	AAATTCATAT	AACCAAAGT	720
TCGAGCCCAA	ACTGGAAGAG	TTCCAGGAGG	TGCAAAGATG	AAAAGAATAA	ATGAAATGAT	780
AAAAGGAGCC	GAATAATGCA	TGACTCCATA	TGGAACCCAG	GCCAAAATAT	CAAGGATGCT	840
ATGCGTGGTT	TTGAGAGAG	GACTAGAAAG	ATTAGAGCCA	TAAAGAATAT	TTTCAAGTGT	900
GGGTAAACA	CGAACCCATA	TGGGTGGACG	CCAGCGTTCT	GGAATAAACC	TACAAGAGTA	960
AAATAAAATT	GCCCAGGTGA	TGATAACAAT	GGCAGGAAAA	AAAATTGGC	GTGTTAAAGG	1020

AACGGTCAAC	GCAATGGCCA	AAAGACAGGC	AATGCCAAAT	TTCCCCCAGA	ATCCAGGAGA	1080
TTCAATGACA	ATACAAGCAA	AAATCAAATT	ACCTGCTAGA	AACACATATT	GCAAATGTGT	1140
CCATGACCAT	TTCGTATTGC	GTAGCAAACG	AAATGTAGGC	ATAGGGTTTA	AGCTTGTTTC	1200
CAACTTGTAT	TGGGATGCTC	GGTTACACGC	AGCAAGGCGC	TTTTTTAAGG	TCGAAAGAGC	1260
AGACATTGCT	TCAAAGAATT	ATCAGAGTAA	AAAAGGGAAG	CGTACGAAAA	AAATTTCGTA	1320
AGGAATTAAC	CGGAAAAC TA	AAGGAAAAAA	AAGGAATTTT	TATGAAGGAA	AGAAAGTAGC	1380
TATTAAATGC	AAGTGTCAAG	CACTTAAAAG	TAGCGATGTA	AAATATTTAA	AAAAAGATGG	1440
ACCGATTAAC	CAATGTTTCA	CTCACAGTTG	CCAGCAATCA	GGGCTATTTT	TTTATTTTTT	1500
TTATAAAATT	GCTAATTATA	TATAATATAA	TTAGTTTATT	AACTTGCTTT	TCCTCAAAAA	1560
ACCAATTTCG	AAGAGGAAC T	TTTGCAGAGG	CAAAAAAGCT	T		1601

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1601 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: mRNA

(iv) ANTI-SENSE: Yes

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGGAAGAUGA	CUUGCAUCAA	AGAUGGAGGA	AGUGGUACUG	GCAGGACGAU	CAAUCAAUUC	60
AGCAGCAGGA	CUAGGUAACG	GCUCAGGUGA	UGAUGAACCC	ACGGACCAUU	CAUGAUCGGU	120

GUUAGCAAGU	UCCAUAUUGU	UAAGACCACU	CAUGAAGGCU	ACUGCAUUAG	GGUUUUGAGU	180
AAAAGAAUCC	CUUCCAAGUA	AGUAUGGGCU	GCCGGUACGA	GCCAAGGAGU	UGCUGGUUUU	240
UUCGGAAAGA	CCAUGACCGU	GGAUAACAAA	CUCGUAUUCC	CAACGAAGGA	UUUUACCAGU	300
UUGCAACUGU	GGGAGGCGUA	GCUUUUGAGC	AAAAACGAAG	CAUAUAAUAG	CUAAACACAU	360
ACCGCCGACC	AAAUUCUACAA	AGUAGUGGUG	GGUAAGGUAC	AUAGUACACC	AGCAAAGCCA	420
UAGAACAUAU	CCAUAAAAGC	AGAAGCGGUA	UCGAGGAAAC	ACAUGCGAAA	GGAAAAGUGC	480
UUCGAGCAUG	GCCCAUCCAG	CGUGAAGAGA	UGGAAAGGCA	CCAAAAACAA	CCGGAGAGUU	540
AGAAAAACCA	UCAGUGUAAA	UGCUGUGGCC	GAAGAGAGCA	UCAAUACGGG	CCAAUCCACC	600
AGGAGAGCCA	CGUACUGCAU	ACGUGGCAGG	UUCUAAACCA	UACAUUUUUU	CAUACCAAGG	660
AGGAGAACAG	GGGAAAGCCA	UUUGGAUAAG	AACACCAAU	AAAUUCAUAU	AACCAAAAGU	720
UCGAGCCCAA	ACUGGAAGAG	UUCGAGGAGG	UGCAAAGAUG	AAAAGAAUAA	AUGAAAUGAU	780
AAAAGGAGCC	GAAUAAUGCA	UGACUCCAUA	UGGAACCCAG	GCCAAAUAU	CAAGGAUGCU	840
AUGCGUGGUU	UUCGAGAGAA	GACUAGAAAG	AUUAGAGCCA	UAAAGAAUAU	UUUCAAGUGU	900
GGGUAAAACA	CGAACCCAUA	UGGGUGGACG	CCAGCGUUCU	GGAAUAAACC	UACAAGAGUA	960
AAAUAAAAUU	GCCCAGGUGA	UGAUAACAAU	GGCAGGAAAA	AAAUUUUGGC	GUGUUAAAGG	1020
AACGGUCAAC	GCAAUGGCCA	AAAGACAGGC	AAUGCCAAU	UUCCCCCAGA	AUCCAGGAGA	1080
UUCAUGACA	AUACAAGCAA	AAAUCAAAUU	ACCUGCUAGA	AACACAUUU	GCAAUGUGU	1140
CCAUGACCAU	UUCGUUUUGC	GUAGCAAACG	AAAUUGUAGG	AUAGGGUUUA	AGCUUGUUUC	1200
CAACUUGUAU	UGGGAUGCUC	GGUACACGC	AGCAAGGCGC	UUUUUUAAAG	UCGAAAGAGC	1260
AGACAUUGCU	UCAAGAAUU	AUCAGAGUAA	AAAAGGGAAG	CGUACGAAA	AAAUUUCGUA	1320
AGGAAUUAAC	CGGAAAACUA	AAGGAAAAAA	AAGGAAUUUU	UAUGAAGGAA	AGAAAGUAGC	1380
UAUUAAAUGC	AAGUGUCAAG	CACUAAAAG	UAGCGAUGUA	AAAUUUUAA	AAAAAGAUGG	1440
ACCGAUUAAC	CAAUGUUCAG	CUCACAGUUG	CCAGCAAUCA	GGGCUAUUUU	UUUAUUUUUU	1500

UUAUAAAAUU GCUAAUUUAU UAUAUAUAA UUAGUUUAUU AACUUGCUUU UCCUCAAAAA 1560
 ACCAAUUCGA GAAAGGAACU UUUGCAGAGG CAAAAAGCU U 1601

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Cys Phe Thr Ser Ser Tyr Phe Pro Asp Asp Arg Arg

5

10

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Cys Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro

5

10

15

Leu Ala Ala Asp

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1553 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

```
TTTTACATAT ATTATTCACC CAATATCATA AAAAAACAA ACTGAATGAT GGCATCTTCT 60
ATTTTGCGTT CCAAATAAT AAAAAACCG TACCAATTAT TCCACTACTA TTTTCTTCTG 120
GAGAAGGCTC CTGGTTCTAC AGTTAGTGAT TTGAATTTTG ATACAAACAT ACAAACGAGT 180
TTACGTAAAT TAAAGCATCA TCATTGGACG GTGGGAGAAA TATTCCATTA TGGGTTTTTG 240
GTTTCCATAC TTTTTTTCGT GTTGTGGTT TTCCCAGCTT CATTTTTTAT AAAATTACCA 300
ATAATCTTAG CATTTGCTAC TTGTTTTTTA ATACCCTTAA CATCACAATT TTTTCTTCCT 360
GCCTTGCCCG TTTTCACTTG GTTGGCATT TATTTTACGT GTGCTAAAAT ACCTCAAGAA 420
TGGAACCAG CTATCACAGT TAAAGTTTTA CCAGCTATGG AAACAATTTT GTACGGCGAT 480
AATTTATCAA ATGTTTTGGC AACCATCACT ACCGGAGTGT TAGATATATT GGCATGGTTA 540
CCATATGGGA TTATTCATTT CAGTTTCCCA TTTGTACTTG CTGCTATTAT ATTTTATTT 600
```

GGGCCACCGA CGGCATTAAG ATCATTTGGA TTTGCCTTTG GTTATATGAA CTTGCTTGGA 660
GTCTTGATTC AAATGGCATT CCCAGCTGCT CCTCCATGGT ACAAAAACCTT GCACGGATTA 720
GAACCAGCTA ATTATTCAAT GCACGGGTCT CCTGGTGGAC TTGGAAGGAT AGATAAATTG 780
TTAGGTGTTG ATATGTATAC CACAGGGTTT TCCAATTCAT CAATCATTTT TGGGGCATTC 840
CCATCGTTAC ATTCAGGATG TTGTATCATG GAAGTGTTAT TTTTGTGTTG GTTGTTTCCA 900
CGATTCAAGT TTGTGTGGGT TACATACGCA TCTTGGCTTT GGTGGAGCAC GATGTATTTG 960
ACCCATCACT ACTTTGTCGA TTTGATTGGT GGAGCCATGC TATCTTTGAC TGTTTTTGAA 1020
TTCACCAAAT ATAAATATTT GCCAAAAAAC AAAGAAGGCC TTTTCTGTCG TTGGTCATAC 1080
ACTGAAATTG AAAAAATCGA TATCCAAGAG ATTGACCCTT TATCATACAA TTATATCCCT 1140
GTCAACAGCA ATGATAATGA AAGCAGATTG TATACGAGAG TGTACCAAGA GCCTCAGGTT 1200
AGTCCCCCAC AGAGAGCTGA AACACCTGAA GCATTTGAGA TGTCAAATTT TTCTAGGTCT 1260
AGACAAAGCT CAAAGACTCA GGTTCATTG AGTAATCTTA CTAACAATGA TCAAGTGCCT 1320
GGAATTAACG AAGAGGATGA AGAAGAAGAA GGCGATGAAA TTTCGTCGAG TACTCCTTCG 1380
GTGTTTGAAG ACGAACCACA GGGTAGCACA TATGCTGCAT CCTCAGCTAC ATCAGTAGAT 1440
GATTTGGATT CAAAAGAAA TTAGTAAAC AGCAGTTTCT ATTAATTTCT TTATTTCTCTC 1500
CTAATTAATG ATTTTATGTT CAATACCTAC ACTATCTGTT TTTAATTTCC TAC 1553

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Met	Met	Ala	Ser	Ser	Ile	Leu	Arg	Ser	Lys	Ile	Ile	Gln	Lys	Pro
1				5					10					15
Tyr	Gln	Leu	Phe	His	Tyr	Tyr	Phe	Leu	Leu	Glu	Lys	Ala	Pro	Gly
				20					25					30
Ser	Thr	Val	Ser	Asp	Leu	Asn	Phe	Asp	Thr	Asn	Ile	Gln	Thr	Ser
				35					40					45
Leu	Arg	Lys	Leu	Lys	His	His	His	Trp	Thr	Val	Gly	Glu	Ile	Phe
				50					55					60
His	Tyr	Gly	Phe	Leu	Val	Ser	Ile	Leu	Phe	Phe	Val	Phe	Val	Val
				65					70					75
Phe	Pro	Ala	Ser	Phe	Phe	Ile	Lys	Leu	Pro	Ile	Ile	Leu	Ala	Phe
				80					85					90
Ala	Thr	Cys	Phe	Leu	Ile	Pro	Leu	Thr	Ser	Gln	Phe	Phe	Leu	Pro
				95					100					105
Ala	Leu	Pro	Val	Phe	Thr	Trp	Leu	Ala	Leu	Tyr	Phe	Thr	Cys	Ala
				110					115					120
Lys	Ile	Pro	Gln	Glu	Trp	Lys	Pro	Ala	Ile	Thr	Val	Lys	Val	Leu
				125					130					135
Pro	Ala	Met	Glu	Thr	Ile	Leu	Tyr	Gly	Asp	Asn	Leu	Ser	Asn	Val
				140					145					150

Leu	Ala	Thr	Ile	Thr	Thr	Gly	Val	Leu	Asp	Ile	Leu	Ala	Trp	Leu
				155					160					165
Pro	Tyr	Gly	Ile	Ile	His	Phe	Ser	Phe	Pro	Phe	Val	Leu	Ala	Ala
				170					175					180
Ile	Ile	Phe	Leu	Phe	Gly	Pro	Pro	Thr	Ala	Leu	Arg	Ser	Phe	Gly
				185					190					195
Phe	Ala	Phe	Gly	Tyr	Met	Asn	Leu	Leu	Gly	Val	Leu	Ile	Gln	Met
				200					205					210
Ala	Phe	Pro	Ala	Ala	Pro	Pro	Trp	Tyr	Lys	Asn	Leu	His	Gly	Leu
				215					220					225
Glu	Pro	Ala	Asn	Tyr	Ser	Met	His	Gly	Ser	Pro	Gly	Gly	Leu	Gly
				230					235					240
Arg	Ile	Asp	Lys	Leu	Leu	Gly	Val	Asp	Met	Tyr	Thr	Thr	Gly	Phe
				245					250					255
Ser	Asn	Ser	Ser	Ile	Ile	Phe	Gly	Ala	Phe	Pro	Ser	Leu	His	Ser
				260					265					270
Gly	Cys	Cys	Ile	Met	Glu	Val	Leu	Phe	Leu	Cys	Trp	Leu	Phe	Pro
				275					280					285
Arg	Phe	Lys	Phe	Val	Trp	Val	Thr	Tyr	Ala	Ser	Trp	Leu	Trp	Trp
				290					295					300
Ser	Thr	Met	Tyr	Leu	Thr	His	His	Tyr	Phe	Val	Asp	Leu	Ile	Gly
				305					310					315
Gly	Ala	Met	Leu	Ser	Leu	Thr	Val	Phe	Glu	Phe	Thr	Lys	Tyr	Lys

320	325	330
Tyr Leu Pro Lys Asn Lys Glu Gly Leu Phe Cys Arg Trp Ser Tyr		
335	340	345
Thr Glu Ile Glu Lys Ile Asp Ile Gln Glu Ile Asp Pro Leu Ser		
350	355	360
Tyr Asn Tyr Ile Pro Val Asn Ser Asn Asp Asn Glu Ser Arg Leu		
365	370	375
Tyr Thr Arg Val Tyr Gln Glu Pro Gln Val Ser Pro Pro Gln Arg		
380	385	390
Ala Glu Thr Pro Glu Ala Phe Glu Met Ser Asn Phe Ser Arg Ser		
395	400	405
Arg Gln Ser Ser Lys Thr Gln Val Pro Leu Ser Asn Leu Thr Asn		
410	415	420
Asn Asp Gln Val Pro Gly Ile Asn Glu Glu Asp Glu Glu Glu Glu		
425	430	435
Gly Asp Glu Ile Ser Ser Ser Thr Pro Ser Val Phe Glu Asp Glu		
440	445	450
Pro Gln Gly Ser Thr Tyr Ala Ala Ser Ser Ala Thr Ser Val Asp		
455	460	465
Asp Leu Asp Ser Lys Arg Asn		
470		

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GACTATTTCA TTATGGGGCC CC 22

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GTTAAC TCGA GAAAGTGCCC ATCAGTGTTC 30

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTTAACGGTA CCTCATCGTT ACACCGTTC 29

(2) INFORMATION FOR SEQ ID NO:40:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 19 bases
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GCTAAACGAC AATCCTGAC 19

(2) INFORMATION FOR SEQ ID NO:41:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 bases

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CGTTGGCCGA TTCATTAATG C 21

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 401 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Met	Ala	Asn	Pro	Phe	Ser	Arg	Trp	Phe	Leu	Ser	Glu	Arg	Pro	Pro
1				5				10					15	
Asn	Cys	His	Val	Ala	Asp	Leu	Glu	Thr	Ser	Leu	Asp	Pro	His	Gln
				20				25					30	
Thr	Leu	Leu	Lys	Val	Gln	Lys	Tyr	Lys	Pro	Ala	Leu	Ser	Asp	Trp
				35				40					45	

Val His Tyr Ile Phe Leu Gly Ser Ile Met Leu Phe Val Phe Ile		
	50	60
Thr Asn Pro Ala Pro Trp Ile Phe Lys Ile Leu Phe Tyr Cys Phe		
	65	75
Leu Gly Thr Leu Phe Ile Ile Pro Ala Thr Ser Gln Phe Phe Phe		
	80	90
Asn Ala Leu Pro Ile Leu Thr Trp Val Ala Leu Tyr Phe Thr Ser		
	95	105
Ser Tyr Phe Pro Asp Asp Arg Arg Pro Pro Ile Thr Val Lys Val		
	110	120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp		
	125	135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp		
	140	150
Leu Pro Tyr Gly Leu Phe His Phe Gly Ala Pro Phe Val Val Ala		
	155	165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr		
	170	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln		
	185	195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly		
	200	210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu		

215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Cys		
230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His		
245	250	255
Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
260	265	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		
275	280	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met		
290	295	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
305	310	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
320	325	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
335	340	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
350	355	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
365	370	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
380	385	390

Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala

395

400

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 401 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Met Ala Asn Pro Phe Ser Arg Trp Phe Leu Ser Glu Arg Pro Pro

1

5

10

15

Asn Cys His Val Ala Asp Leu Glu Thr Ser Leu Asp Pro His Gln

20

25

30

Thr Leu Leu Lys Val Gln Lys Tyr Lys Pro Ala Leu Ser Asp Trp

35

40

45

Val His Tyr Ile Phe Leu Gly Ser Ile Met Leu Phe Val Phe Ile

50

55

60

Thr Asn Pro Ala Pro Trp Ile Phe Lys Ile Leu Phe Tyr Cys Phe

65

70

75

Leu Gly Thr Leu Phe Ile Ile Pro Ala Thr Ser Gln Phe Phe Phe

	80	85	90
Asn Ala Leu Pro Ile Leu Thr Trp Val Ala Leu Tyr Phe Thr Ser			
	95	100	105
Ser Tyr Phe Pro Asp Asp Arg Arg Pro Pro Ile Thr Val Lys Val			
	110	115	120
Leu Pro Ala Val Glu Thr Ile Leu Tyr Gly Asp Asn Leu Ser Asp			
	125	130	135
Ile Leu Ala Thr Ser Thr Asn Ser Phe Leu Asp Ile Leu Ala Trp			
	140	145	150
Leu Pro Tyr Gly Leu Phe His Tyr Gly Ala Pro Phe Val Val Ala			
	155	160	165
Ala Ile Leu Phe Val Phe Gly Pro Pro Thr Val Leu Gln Gly Tyr			
	170	175	180
Ala Phe Ala Phe Gly Tyr Met Asn Leu Phe Gly Val Ile Met Gln			
	185	190	195
Asn Val Phe Pro Ala Ala Pro Pro Trp Tyr Lys Ile Leu Tyr Gly			
	200	205	210
Leu Gln Ser Ala Asn Tyr Asp Met His Gly Ser Pro Gly Gly Leu			
	215	220	225
Ala Arg Ile Asp Lys Leu Leu Gly Ile Asn Met Tyr Thr Thr Cys			
	230	235	240
Phe Ser Asn Ser Ser Val Ile Phe Gly Ala Phe Pro Ser Leu His			
	245	250	255

Ser Gly Cys Ala Thr Met Glu Ala Leu Phe Phe Cys Tyr Cys Phe		
	260	270
Pro Lys Leu Lys Pro Leu Phe Ile Ala Tyr Val Cys Trp Leu Trp		
	275	285
Trp Ser Thr Met Tyr Leu Thr His His Tyr Phe Val Asp Leu Met		
	290	300
Ala Gly Ser Val Leu Ser Tyr Val Ile Phe Gln Tyr Thr Lys Tyr		
	305	315
Thr His Leu Pro Ile Val Asp Thr Ser Leu Phe Cys Arg Trp Ser		
	320	330
Tyr Thr Ser Ile Glu Lys Tyr Asp Ile Ser Lys Ser Asp Pro Leu		
	335	345
Ala Ala Asp Ser Asn Asp Ile Glu Ser Val Pro Leu Ser Asn Leu		
	350	360
Glu Leu Asp Phe Asp Leu Asn Met Thr Asp Glu Pro Ser Val Ser		
	365	375
Pro Ser Leu Phe Asp Gly Ser Thr Ser Val Ser Arg Ser Ser Ala		
	380	390
Thr Ser Ile Thr Ser Leu Gly Val Lys Arg Ala		
	395	400

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```

ATGGCAAACC CTTTTTCGAG ATGGTTTCTA TCAGAGAGAC CTCCAAACTG CCATGTAGCC   60
GATTTAGAAA CAAGTTTAGA TCCCCATCAA ACGTTGTTGA AGGTGCAAAA ATACAAACCC  120
GCTTTAAGCG ACTGGGTGCA TTACATCTTC TTGGGATCCA TCATGCTGTT TGTGTTCATT  180
ACTAATCCCG CACCTTGGAT CTTCAAGATC CTTTTTTATT GTTTCTTGGG CACTTTATTC  240
ATCATTCAG CTACGTCACA GTTTTTCTTC AATGCCTTGC CCATCCTAAC ATGGGTGGCG  300
CTGTATTTCA CTTTCATCGTA CTTTCCAGAT GACCGCAGGC CTCCTATTAC TGTCAAAGTG  360
TTACCAGCGG TGGAAACAAT TTTATACGGC GACAATTTAA GTGATATTCT TGCAACATCG  420
ACGAATTCCT TTTTGGACAT TTTAGCATGG TTACCGTACG GACTATTTCA TTTTGGGGCC  480
CCATTTGTCG TTGCTGCCAT CTTATTCGTA TTTGGTCCAC CAACTGTTTT GCAAGGTTAT  540
GCTTTTGCAT TTGGTTATAT GAACCTGTTT GGTGTTATCA TGCAAAATGT CTTTCCAGCC  600
GCTCCCCCAT GGTATAAAAT TCTCTATGGA TTGCAATCAG CCAACTATGA TATGCATGGC  660
TCGCCTGGTG GATTAGCTAG AATTGATAAG CTACTCGGTA TTAATATGTA TACTACATGT  720
TTTTCAAATT CCTCCGTCAT TTTCGGTGCT TTTCCTTCAC TGCATTCCGG GTGTGCTACT  780
ATGGAAGCCC TGTTTTTCTG TTATTGTTTT CCAAATGA AGCCCTTGTT TATTGCTTAT  840
GTTTGCTGGT TATGGTGGTC AACTATGTAT CTGACACACC ATTATTTTGT AGACCTTATG  900
GCAGGTTCTG TGCTGTCATA CGTTATTTTC CAGTACACAA AGTACACACA TTTACCAATT  960

```

GTAGATACAT CTCTTTTTTG CAGATGGTCA TACACTTCAA TTGAGAAATA CGATATATCA 1020
AAGAGTGATC CATTGGCTGC AGATTCAAAC GATATCGAAA GTGTCCCTTT GTCCAACCTG 1080
GAACTTGACT TTGATCTTAA TATGACTGAT GAACCCAGTG TAAGCCCTTC GTTATTTGAT 1140
GGATCTACTT CTGTTTCTCG TTCGTCCGCC ACGTCTATAA CGTCACTAGG TGTAAAGAGG 1200
GCTTAA 1206

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1206 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATGGCAAACC CTTTTTCGAG ATGGTTTCTA TCAGAGAGAC CTCCAAACTG CCATGTAGCC 60
GATTTAGAAA CAAGTTTAGA TCCCCATCAA ACGTTGTTGA AGGTGCAAAA ATACAAACCC 120
GCTTTAAGCG ACTGGGTGCA TTACATCTTC TTGGGATCCA TCATGCTGTT TGTGTTTCATT 180
ACTAATCCCG CACCTTGGAT CTTCAAGATC CTTTTTTATT GTTCTTGGG CACTTTATTC 240
ATCATTCAG CTACGTCACA GTTTTTCTTC AATGCCTTGC CCATCCTAAC ATGGGTGGCG 300
CTGTATTTCA CTTTCATCGTA CTTTCCAGAT GACCGCAGGC CTCCTATTAC TGTCAAAGTG 360
TTACCAGCGG TGGAAACAAT TTTATACGGC GACAATTTAA GTGATATTCT TGCAACATCG 420
ACGAATTCCT TTTTGGACAT TTTAGCATGG TTACCGTACG GACTATTTCA TTATGGGGCC 480

CCATTTGTCG	TTGCTGCCAT	CTTATTCGTA	TTTGGTCCAC	CAACTGTTTT	GCAAGGTTAT	540
GCTTTTGCAT	TTGGTTATAT	GAACCTGTTT	GGTGTATCA	TGCAAAATGT	CTTTCCAGCC	600
GCTCCCCCAT	GGTATAAAAT	TCTCTATGGA	TTGCAATCAG	CCAACATGA	TATGCATGGC	660
TCGCCTGGTG	GATTAGCTAG	AATTGATAAG	CTACTCGGTA	TTAATATGTA	TACTACATGT	720
TTTTCAAATT	CCTCCGTCAT	TTTCGGTGCT	TTTCCTTCAC	TGCATTCCGG	GTGTGCTACT	780
ATGGAAGCCC	TGTTTTTCTG	TTATTGTTTT	CCAAAATTGA	AGCCCTTGTT	TATTGCTTAT	840
GTTTGCTGGT	TATGGTGGTC	AACTATGTAT	CTGACACACC	ATTATTTTGT	AGACCTTATG	900
GCAGGTTCTG	TGCTGTCATA	CGTTATTTTC	CAGTACACAA	AGTACACACA	TTTACCAATT	960
GTAGATACAT	CTCTTTTTTG	CAGATGGTCA	TACACTTCAA	TTGAGAAATA	CGATATATCA	1020
AAGAGTGATC	CATTGGCTGC	AGATTCAAAC	GATATCGAAA	GTGTCCCTTT	GTCCAACCTG	1080
GAACCTGACT	TTGATCTTAA	TATGACTGAT	GAACCCAGTG	TAAGCCCTTC	GTTATTTGAT	1140
GGATCTACTT	CTGTTTCTCG	TTCGTCCGCC	ACGTCTATAA	CGTCACTAGG	TGTAAAGAGG	1200
GCTTAA						1206

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1206 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATGGCAAACC	CTTTTTTCGAG	ATGGTTTCTA	TCAGAGAGAC	CTCCAAACTG	CCATGTAGCC	60
GATTTAGAAA	CAAGTTTAGA	TCCCCATCAA	ACGTTGTTGA	AGGTGCAAAA	ATACAAACCC	120
GCTTTAAGCG	ACTGGGTGCA	TTACATCTTC	TTGGGATCCA	TCATGCTGTT	TGTGTTCAATT	180
ACTAATCCCG	CACCTTGGAT	CTTCAAGATC	CTTTTTTATT	GTTTCTTGGG	CACCTTATTC	240
ATCATTCAG	CTACGTCACA	GTTTTTCTTC	AATGCCTTGC	CCATCCTAAC	ATGGGTGGCG	300
CTGTATTTCA	CTTCATCGTA	CTTCCAGAT	GACCGCAGGC	CTCCTATTAC	TGTCAAAGTG	360
TTACCAGCGG	TGGAAACAAT	TTTATACGGC	GACAATTTAA	GTGATATTCT	TGCAACATCG	420
ACGAATTCCT	TTTTGGACAT	TTTAGCATGG	TTACCGTACG	GACTATTTCA	TTATGGGGCC	480
CCATTGTGCG	TTGCTGCCAT	CTTATTCGTA	TTTGGTCCAC	CAACTGTTTT	GCAAGGTTAT	540
GCTTTTGCA	TTGGTTATAT	GAACCTGTTT	GGTGTATCA	TGCAAAATGT	CTTCCAGCC	600
GCTCCCCCAT	GGTATAAAAT	TCTCTATGGA	TTGCAATCAG	CCAACATGA	TATGCATGGC	660
TCGCCTGGTG	GATTAGCTAG	AATTGATAAG	CTACTCGGTA	TTAATATGTA	TACTACAGCT	720
TTTTCAAATT	CCTCCGTCAT	TTTCGGTGCT	TTTCCTTCAC	TGCATTCCGG	GTGTGCTACT	780
ATGGAAGCCC	TGTTTTTCTG	TTATTGTTTT	CCAAAATTGA	AGCCCTTGTT	TATTGCTTAT	840
GTTTGCTGGT	TATGGTGGTC	AACTATGTAT	CTGACACACC	ATTATTTTGT	AGACCTTATG	900
GCAGGTTCTG	TGCTGTCATA	CGTTATTTTC	CAGTACACAA	AGTACACACA	TTTACCAATT	960
GTAGATACAT	CTCTTTTTTG	CAGATGGTCA	TACACTTCAA	TTGAGAAATA	CGATATATCA	1020
AAGAGTGATC	CATTGGCTGC	AGATTCAAAC	GATATCGAAA	GTGTCCCTTT	GTCCAACCTG	1080
GAACCTGACT	TTGATCTTAA	TATGACTGAT	GAACCCAGTG	TAAGCCCTTC	GTTATTTGAT	1140
GGATCTACTT	CTGTTTCTCG	TTCGTCCGCC	ACGTCTATAA	CGTCACTAGG	TGTAAAGAGG	1200
GCTTAA						1206

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1206 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

```
ATGGCAAACC CTTTTTCGAG ATGGTTTCTA TCAGAGAGAC CTCCAAACTG CCATGTAGCC 60
GATTTAGAAA CAAGTTTAGA TCCCCATCAA ACGTTGTTGA AGGTGCAAAA ATACAAACCC 120
GCTTTAAGCG ACTGGGTGCA TTACATCTTC TTGGGATCCA TCATGCTGTT TGTGTTCATT 180
ACTAATCCCG CACCTTGGAT CTTCAAGATC CTTTTTTATT GTTTCTTGGG CACTTTATTC 240
ATCATTCAG CTACGTCACA GTTTTTCTTC AATGCCTTGC CCATCCTAAC ATGGGTGGCG 300
CTGTATTTCA CTTTCATCGTA CTTTCCAGAT GACCGCAGGC CTCCTATTAC TGTCAAAGTG 360
TTACCAGCGG TGGAAACAAT TTTATACGGC GACAATTTAA GTGATATTCT TGCAACATCG 420
ACGAATTCCT TTTTGGACAT TTTAGCATGG TTACCGTACG GACTATTTCA TTTTGGGGCC 480
CCATTTGTCG TTGCTGCCAT CTTATTCGTA TTTGGTCCAC CAACTGTTTT GCAAGGTTAT 540
GCTTTTGCAT TTGGTTATAT GAACCTGTTT GGTGTTATCA TGCAAAATGT CTTTCCAGCC 600
GCTCCCCCAT GGTATAAAAT TCTCTATGGA TTGCAATCAG CCAACTATGA TATGCATGGC 660
TCGCCTGGTG GATTAGCTAG AATTGATAAG CTACTCGGTA TTAATATGTA TACTACAGCT 720
TTTTCAAATT CCTCCGTCAT TTTCGGTGCT TTTCCTTCAC TGCATTCCGG GTGTGCTACT 780
ATGGAAGCCC TGTTTTTCTG TTATTGTTTT CCAAATGA AGCCCTTGTT TATTGCTTAT 840
GTTTGCTGGT TATGGTGGTC AACTATGTAT CTGACACACC ATTATTTTGT AGACCTTATG 900
```

GCAGGTTCTG TGCTGTCATA CGTTATTTTC CAGTACACAA AGTACACACA TTTACCAATT 960
 GTAGATACAT CTCTTTTTTG CAGATGGTCA TACACTTCAA TTGAGAAATA CGATATATCA 1020
 AAGAGTGATC CATTGGCTGC AGATTCAAAC GATATCGAAA GTGTCCCTTT GTCCAACCTG 1080
 GAACTTGACT TTGATCTTAA TATGACTGAT GAACCCAGTG TAAGCCCTTC GTTATTTGAT 1140
 GGATCTACTT CTGTTTCTCG TTCGTCCGCC ACGTCTATAA CGTCACTAGG TGTAAGAGG 1200
 GCTTAA 1206

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AATATGTATA CTACATGTTT TTCAAATTCC 30

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GTAACTCGA GAAAGTGCCC ATCAGTGTTT 30

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid (synthetic DNA)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GTTAACGGTA CCAGAGGAAA GAATAACGC 29